

PD - 7/14/2002
Ran 4 + W/O

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 320.034 Seconds
(without alignments)
650.171 Million cell updates/sec

Title: US-10-671-242-43
Perfect score: 2850
Sequence: 1 MAGPPALPPPTAAATAA.....HIKVLQGHFDDPDGFLG 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2850	100.0	538	4	AAB82806	Human low
2	2594	91.0	550	4	AAB82807	Rabbit lo
3	1445.5	50.7	317	4	AAB82798	Rabbit lo
4	1439.5	50.5	317	2	AAW49038	Rabbit lo
5	1260	44.2	241	3	AAW58957	Breast an
6	1191.5	41.8	252	4	AAB82800	Rabbit lo
7	1170	41.1	217	2	AAW49041	Human low
8	1170	41.1	217	4	AAB82803	Human low
9	1086.5	38.1	232	4	AAB82799	Rabbit lo
10	708	24.8	132	7	AD59158	Human Pro
11	478	16.8	316	4	AAU31793	Novel hum
12	340	11.9	801	8	ADQ97686	Mouse can
13	336	11.8	783	2	AAW37151	Mouse neu
14	336	11.8	787	2	AAW37152	Mouse neu
15	336	11.8	802	2	AAW37153	Mouse neu
16	336	11.8	802	4	AAU09139	Mammalian
17	324	11.4	397	8	ADQ97684	Mouse can
18	323	11.3	639	8	ADN21309	Bacterial
19	319.5	11.2	783	7	ABO69289	Pseudomon
20	312	10.9	61	3	AAW34365	Human sec
21	311.5	10.9	1151	4	ABB61598	Drosophila
22	311	10.9	416	5	ABG70108	Human pre
23	311	10.9	505	2	AAW46889	Human Neu
24	311	10.9	505	2	AAW53808	N-Wiskott
25	311	10.9	505	4	AAG67338	Amino aci

ALIGNMENTS

RESULT 1
AAB82806

ID AAB82806 standard; protein; 538 AA.

XX AC AAB82806;

DT 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 2 (LBP-2).

KW Low density lipoprotein binding protein 2; LBP-2; LDL; human; atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

OS Homo sapiens.

PN WO200164874-A2.

XX 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US006356.

PR 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAH26499.

XX New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence was deduced from the coding region of isolated genomic DNA (see AAH26499). It differs from the sequence predicted from an LBP-2 cDNA clone (see AAH2803) by the presence of an additional 321 amino acids at the N-terminus (the cDNA clone is 5' truncated). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529-538 (see AAH2803-12) of the present sequence are claimed. Methods of

CC determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX
 SQ Sequence 538 AA;

Query Match 100.0%; Score 2850; DB 4; Length 538;
 Best Local Similarity 100.0%; Pred. No. 3.5e-165;
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPPALPPPTAAATTAAASSSSAASPHYQEWILDTIDSLRSKARPDLERICRMVR 60
 DB 1 MAGPPALPPPTAAATTAAASSSSAASPHYQEWILDTIDSLRSKARPDLERICRMVR 60
 QY 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120
 DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120
 QY 121 AAAAAAPPPTTAPPPPPAPVAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180
 DB 121 AAAAAAPPPTTAPPPPPAPVAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180
 QY 181 PPAPAPPAVAPPAGPRRAPPPAARVAREPLPPPPPPAPPOQQQPPPPQPPPPGGAV 240
 DB 181 PPAPAPPAVAPPAGPRRAPPPAARVAREPLPPPPPPAPPOQQQPPPPQPPPPGGAV 240
 QY 241 RAGGAARPVSLREVVRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGALALPRG 300
 DB 241 RAGGAARPVSLREVVRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGALALPRG 300
 QY 301 DRPGRAAPPAASARPSKRGGEERVLKEEEDDDDEDEEDDVSGSEVPSPDRPAGAQ 360
 DB 301 DRPGRAAPPAASARPSKRGGEERVLKEEEDDDDEDEEDDVSGSEVPSPDRPAGAQ 360
 QY 361 HHOLNCGERGQSAKERVKWTCPGHQGDGGRGAPGSGTRQVFSMAANKEGGTASVA 420
 DB 361 HHOLNCGERGQSAKERVKWTCPGHQGDGGRGAPGSGTRQVFSMAANKEGGTASVA 420
 QY 421 TGPDSPPVPLPGPKALPGADGTPGCPGPKRKEKSDPVWVTMDVVEYFTGAPPEQA 480
 DB 421 TGPDSPPVPLPGPKALPGADGTPGCPGPKRKEKSDPVWVTMDVVEYFTGAPPEQA 480
 QY 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538
 DB 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538

RESULT 2

AAB82807
 ID AAB82807 standard; protein; 550 AA.

XX AAB82807;

XX 12-NOV-2001 (first entry)

XX Rabbit low density lipoprotein binding protein 2 (LBP-2).

DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiatherosclerotic; therapy; diagnosis; vaccine.

XX Oryctolagus cuniculus.

OS WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

PA (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

DR N-PSDB; AAH26500.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(k); Fig 2a; 143pp; English.

XX The present sequence is that of a full-length sequence of novel rabbit
 CC low density lipoprotein binding protein 2 (LBP-2). The amino acid
 CC sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit
 CC LBP-2 is an example of claimed polypeptides of the invention, termed
 CC LBP-2, that are capable of binding to native and methylated low density
 CC lipoproteins. Also claimed are biologically active fragments and
 CC analogues of LBPs, polynucleotides encoding LBPs, as well as expression
 CC vectors, cells and methods of producing the LBPs. Polypeptides having
 CC amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-
 CC 18) of the present sequence are claimed. Methods for determining if an
 CC animal is at risk for atherosclerosis, methods for evaluating an agent
 CC for use in treating atherosclerosis, and methods for treating a cell
 CC having an abnormality in structure or metabolism of LBP are also claimed,
 CC as are pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions

XX Sequence 550 AA;

Query Match 91.0%; Score 2594; DB 4; Length 550;
 Best Local Similarity 91.1%; Pred. No. 1.3e-149;
 Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPTAAATT-AAAASSAASPHYQEWILDTIDSLRSKARPDLERICRMVR 59

DB 1 MAGPPALPPPTAAATTAAASSSSAASPHYQEWILDTIDSLRSKARPDLERICRMVR 60

QY 60 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 119

DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120

QY 120 AAAAAAPPPTTAPPPPPAPV-AAAAAPARAP-AAAAAATAPSPGPAQPGPRAQRAAP 176

DB 121 --AAAAAPPPTTAPPPPPAPVAAAPARAPRAAAAAAATAPSPGPAQPGPRAQRAAP 178

QY 177 LAAPPAPAAAPVAPPAGPRRAPPP--AAVAAREPPLPPPPPPPPPPPPPPPPPP 229

DB 179 LAAPPAPAAAPVAPPAGPRRAPPPAAVAAREPPLPPPPPPPPPPPPPPPPPP 238

QY 230 QPQPPPEGGAAGGAAARVPSLRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLER 289

DB 239 QPQPPPEGGAAGGAAARVPSLRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLER 298

QY 290 TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLKEEEDDDDEDEEDD--VSEG 347

DB 299 TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLKEEEDDDDEDEEDD--VSEG 358

QY 348 SEVPESDRPAGAQHHQLN-GERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTQVFS 406

DB 359 SEVPESDRPAGAQHHQLNGERGTPQAKERAKESWLCGPHQGEGRGPAAGSGTQVFS 418

QY 407 MAANKEGGTASVATGPDSPSPVPLPGKPAALPGADGTPGCPGPKRKEKPSDPVEVTMD 466

DB 419 MAALSKEGGSASTTGPDSPSPVPLPGKPAALPGADGTPGCPGPKRKEKPSDPVEVTMD 478

QY 467 VVEYFTGAPPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 526

DB 479 VVEYFTGAPPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 538

QY 527 HFEDDDPDGFLG 538

DB 539 HFEDDDPDGFLG 550

Db	131	SDR	PAGQHQLNGERCPQTAKERAKENSLCGPHPGQEEGRGPAAGSGTRQVFSMAALS	191					
Qy	412	KEG	TASVATGPDSPSPVPLPPGKPALPGADGTPFGCPGGRKEKPSDPVEWTVMDVVEYF	471					
Db	191	KEG	GSASSTTGPDSPSPVPLPPGKPALPGADGTPFGCPGGRKEKPADPVEWTVMDVVEYF	250					
Qy	472	TEAG	FPQATAFQOEIDGKSLLMQRTDVLGTLSIRLGPAKLIYEHHIKVLOQGHFEDD	531					
Db	251	TEAG	FPQATAFQOEIDGKSLLMQRTDVLGTLSIRLGPAKLIYEHHIKVLOQGHFEDD	310					
Qy	532	DPD	GFLG 538						
Db	311	DPE	GFLG 317						
RESULT 4									
ID	AAW49038	standard; protein; 317 AA.							
XX	AAW49038;								
XX	09-NOV-1998	(first entry)							
XX	Rabbit low density lipoprotein binding protein LBP-2.								
DE									
KW	Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;								
KW	receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.								
OS	Oryctolagus cuniculus.								
XX									
FH	Key	Location/Qualifiers							
FT	Misc-difference 10	/note= "encoded by TAG"							
FT	Protein	66..317	/note= "Claim 1"						
FT	Protein	86..317	/note= "Claim 1"						
FT	Peptide	105..132	/note= "Claim 2"						
FT	Peptide	105..120	/note= "Claim 2"						
FT	Peptide	121..132	/note= "Claim 2"						
FT	Peptide	211..220	/note= "Claim 2"						
XX	WO9823282-A1.								
XX	ED	04-JUN-1998							
XX	26-NOV-1997;	97WO-US021857.							
XX	27-NOV-1996;	96US-0031930P.							
PR	03-JUN-1997;	97US-0048547P.							
XX	(BOST-) BOSTON HEART FOUND INC.								
XX	Lees AM, Lees RS, Law SW, Arjona AA;								
XX	WPI; 1998-322455/28.								
DR	N-PSDE; AAV32835.								
XX	Nucleic acid encoding low density lipoprotein binding proteins and								
PT	related vectors - transformed cells, proteins, and modulators of binding,								
PT	useful for treatment and diagnosis of atherosclerosis and for identifying								
PT	subjects at risk.								
XX	Claim 1; Fig 2; 47pp; English.								
XX	This polypeptide, comprises novel rabbit low density lipoprotein (LDL)								
CC	binding protein LBP-2 that is capable of binding both native and methyl								
CC	LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA								
CC	(see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and								


```

XX PN WO200164874-A2.
XX PD 07-SEP-2001.
XX PF 28-FEB-2001; 2001WO-US006356.
XX PR 02-MAR-2000; 2000US-00517849.
XX PR 14-JUL-2000; 2000US-00616289.
XX PA (BOST-) BOSTON HEART FOUND INC.
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX DR WPI; 2001-565505/63.
XX DR N-PSDB; AAH26490.
XX PT New isolated low density lipoprotein binding polypeptide for treating,
XX PS diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX PF Claim 13(d); Fig 4; 143pp; English.
XX CC The present sequence is that of a partial sequence of novel rabbit low
XX CC density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is
XX CC deduced from an isolated cDNA clone (see AAH26490). A full-length
XX CC sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an
XX CC example of claimed polypeptides of the invention, termed LBPs, that are
XX CC capable of binding to native and methylated low density lipoproteins.
XX CC Also claimed are biologically active fragments and analogues of LBPs,
XX CC polynucleotides encoding LBPs, as well as expression vectors, cells and
XX CC methods of producing the LBPs. Methods of determining if an animal is at
XX CC risk for atherosclerosis, methods for evaluating an agent for use in
XX CC treating atherosclerosis, and methods for treating a cell having an
XX CC abnormality in structure or metabolism of LBP are also claimed, as are
XX CC pharmaceutical compositions comprising an LBP polypeptide or nucleic
XX CC acid, and vaccine compositions
XX SQ Sequence 252 AA;

Query Match 41.8%; Score 1191.5; DB 4; Length 252;
Best Local Similarity 88.5%; Pred. No. 1e-64;
Matches 223; Conservative 17; Mismatches 9; Indels 3; Gaps 2;

QY 290 TRIGALALPRGDRPGRAPPAASARPSKRGSEVLEKEEEDDDDEDEDD--VSEG 347
Db 1 TRIGALALPRGDRPGRAPPAASARARANKRAGEERVELEKEEEDDDDDDDVVSEG 60

QY 348 SEVPSDRPAGAHQLN-CERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFS 406
Db 61 SEVPSDRPAGAHQLNGERGPGTAKERAKESWLCGPHGQEGRGPAAGSGTRQVFS 120

QY 407 MAAMNKEGGTASVATGPDSPSPVPLPGKALPGADGTPGCPGRKEKSDPVWTVMD 466
Db 121 MAALSKEGGASSTTGPDSPSPVPLPGKALPGADGTPGCPGRKEKSDPVWTVMD 180

QY 467 VVEYFTAGPPEQATAFQEQIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIVLQQG 526
Db 181 VVEYFTAGPPEQATAFQEQIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIVLQQG 240

QY 527 HFEDDDPDGFLG 538
Db 241 HFEDDDPDGFLG 252

RESULT 7
AAW49041
ID AAW49041 standard; protein; 217 AA.
XX AC AAW49041;
XX XX
XX DT 09-NOV-1998 (first entry)
XX DE Human low density lipoprotein binding protein LBP-2.

Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
receptor; human; atherosclerosis; diagnosis; therapy; vaccine.
Homo sapiens.
Key Location/Qualifiers
FT Peptide 8..33
FT Peptide /note= "Claim 2"
FT Peptide 8..22
FT Peptide /note= "Claim 2"
FT Peptide 23..33
FT Peptide /note= "Claim 2"
FT Peptide 208..217
FT Peptide /note= "Claim 2"
PN WO9823282-A1.
XX 04-JUN-1998.
XX 26-NOV-1997; 97WO-US021857.
XX 27-NOV-1996; 96US-0031930P.
XX 03-JUN-1997; 97US-0048547P.
XX (BOST-) BOSTON HEART FOUND INC.
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX DR WPI; 1998-322455/28.
XX DR N-PSDB; AAV32838.
XX Nucleic acid encoding low density lipoprotein binding proteins and
XX related vectors - transformed cells, proteins, and modulators of binding,
XX useful for treatment and diagnosis of atherosclerosis and for identifying
XX subjects at risk.
XX Claim 1; Fig 7; 47pp; English.
XX This polypeptide comprises novel human low density lipoprotein (LDL)
XX binding protein LBP-2 that is capable of binding both native and methyl
XX LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
XX AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human
XX LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP
XX metabolism or structure is diagnostic of a risk for atherosclerosis. The
XX invention provides methods for determining if an animal is at risk for
XX atherosclerosis (e.g. for prenatal screening); methods for treating
XX atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
XX bind LDL and thereby prevent formation of atherosclerotic plaque; and
XX methods for treating a cell having an abnormality in LBP structure or
XX metabolism. Pharmaceutical and vaccine compositions are also provided, as
XX well as recombinant vectors and host cells used to produce recombinant
XX LBP
XX Sequence 217 AA;

Query Match 41.1%; Score 1170; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e-63;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 EERVLEKEEEDDDDEDEDDVSEGSEVPSDRPAGAHQLNGERGQSAKERVKEWT 381
Db 1 EERVLEKEEEDDDDEDEDDVSEGSEVPSDRPAGAHQLNGERGQSAKERVKEWT 60

QY 382 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKALPGA 441
Db 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKALPGA 120

QY 442 DGTTPGCPGRKEKSDPVWTVMDVVEYFTAGPPEQATAFQEQIDGKSLLMQRTDV 501
Db 121 DGTTPGCPGRKEKSDPVWTVMDVVEYFTAGPPEQATAFQEQIDGKSLLMQRTDV 180

QY 502 LTGLSIRLGPALKIYEHKIVLQQGHFEDDDPDGFLG 538

```

Db 181 LTGLSIRLGPALKIYEHKIKVLQGGHFEDDDPGFLG 217

RESULT 8
 AAB82803
 ID AAB82803 standard; protein; 217 AA.
 AC AAB82803;
 XX
 XX 12-NOV-2001 (first entry)
 XX
 DE Human low density lipoprotein binding protein 2 (LBP-2).
 XX
 XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US006356.
 XX
 XX 02-MAR-2000; 2000US-00517849.
 PR
 PR 14-JUL-2000; 2000US-00616289.
 XX
 XX (BOST-) BOSTON HEART FOUND INC.
 PA
 XX
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 PI
 XX WPI; 2001-565505/63.
 DR
 DR N-PSDB; AAH26494.
 XX
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 PT
 PS Claim 13(g); Fig 7B; 143pp; English.
 XX

The present sequence is that of the N-terminal portion of novel human low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is deduced from an isolated partial cDNA clone (see AAH26494). A full-length sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are claimed. Pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine compositions, are also claimed

SQ Sequence 217 AA;

Query Match 41.1%; Score 1170; DB 4; Length 217;
 Best Local Similarity 100.0%; Pred. No. 1.8e-63;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 EERVLSKEEEDDEDEDEDDVSEGVPESDRPAQAQHQLNGRGPQSAKERVKEWT 381
 Db 1 EERVLSKEEEDDEDEDEDDVSEGVPESDRPAQAQHQLNGRGPQSAKERVKEWT 60
 QY 382 PCGPHQOQDEGRGPAFCGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 441
 Db 61 PCGPHQOQDEGRGPAFCGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
 QY 442 DGTFFCGPKRKEKPSDPVWNTVMDVVEYFTGAFPEQATAFQEQIDGKSLLLMORTDV 501
 Db 121 DGTFFCGPKRKEKPSDPVWNTVMDVVEYFTGAFPEQATAFQEQIDGKSLLLMORTDV 180
 QY 502 LTGLSIRLGPALKIYEHKIKVLQGGHFEDDDPGFLG 538

Db 181 LTGLSIRLGPALKIYEHKIKVLQGGHFEDDDPGFLG 217

RESULT 9
 AAB82799
 ID AAB82799 standard; protein; 232 AA.
 AC AAB82799;
 XX
 XX 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
 XX
 XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US006356.
 XX
 XX 02-MAR-2000; 2000US-00517849.
 PR
 PR 14-JUL-2000; 2000US-00616289.
 XX
 XX (BOST-) BOSTON HEART FOUND INC.
 PA
 XX
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 PI
 XX WPI; 2001-565505/63.
 DR
 DR N-PSDB; AAH26489.
 XX
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 PT
 PS Claim 13(c); Fig 3; 143pp; English.
 XX

The present sequence is that of a partial sequence of novel rabbit low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is deduced from an isolated cDNA clone (see AAH26489). Full-length rabbit LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed LBP polypeptides of the invention, termed LBPs, that are capable of binding to native and methylated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine compositions

SQ Sequence 232 AA;

Query Match 38.1%; Score 1086.5; DB 4; Length 232;
 Best Local Similarity 87.5%; Pred. No. 2.3e-58;
 Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;

QY 310 ASARSRSRKGGFERVLEKEEEDDEDEDEDD--VSEGVPESDRPAQAQHQLN-G 366
 Db 1 ASARAARNKRAEERVLEKEEEDDEDEDEDDDDDDVVSEGVPESDRPAQAQHQLNGG 60
 QY 367 ERGPQSAKERVKEWTPCGPHQOQDEGRGPAFCGSGTRQVFSMAAMNKEGGTASVATGPDSP 426
 Db 61 ERGPQTAKEAKESWLSLGGHPHCGEEGRGPAFCGSGTRQVFSMAALSKEGGSASSTTGPDS 120
 QY 427 SPVPLPPGKPALPGADGTPFGCPGPKRKEKPSDPVWNTVMDVVEYFTGAFPEQATAFQEQ 486
 Db 121 SPVPLPPGKPALPGADGTPFGCPGPKRKEKPSDPVWNTVMDVVEYFTGAFPEQATAFQEQ 180

Qy 487 EIDGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEDEDDPDGFLG 538
Db 181 EIDGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEDEDDPDGFLG 232

RESULT 10
ADE59158
ID ADE59158 standard; protein; 132 AA.
XX
AC ADE59158;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human Protein XP_031299, SEQ ID NO 5049.
DE Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346392P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-269312/26.
DR GENBANK; XP_031299.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 132 AA;

Query Match 24.8%; Score 708; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.3e-35;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 MAAMNKEGGTASVATGPDSPVPLPGKPALPGADGTPFGCPGRKEKPSDPVETVMD 466
Db 1 MAAMNKEGGTASVATGPDSPVPLPGKPALPGADGTPFGCPGRKEKPSDPVETVMD 60

Qy 467 VVEYFTGAGPPEQATFOEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQOQ 526
Db 61 VVEYFTGAGPPEQATFOEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQOQ 120

Qy 527 HFEDDDPDGFLG 538
Db 121 HFEDDDPDGFLG 132

RESULT 11
AAU31793
ID AAU31793 standard; protein; 316 AA.
XX
AC AAU31793;
XX
XX 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #2284.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
XX
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
DR
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
XX Claim 20; Page 509; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX
XX Sequence 316 AA;

Query Match 16.8%; Score 478; DB 4; Length 316;
Best Local Similarity 38.5%; Pred. No. 3e-21;
Matches 136; Conservative 9; Mismatches 84; Indels 124; Gaps 14;

QY 235 PEGAVRAGGAARFVSLREYVYLGSGGAGGRLTRGRVQGLLEEEAARGRLERLGA 294
DB 2 PRAGAAGGG-----GGRPSGAHPSR-----SA 25
QY 295 LALPRGRDRGRAPPAASARSRKRGGEERVLKKEEEDD-DEDEEDDVSGSV-PE 352
DB 26 CGCPAGTRPGRAPPAASARSRKRGGEERVLKKEEEDDDEKXRXRRSMCQRSEVAPS 85
QY 353 SDRPAGAHQHLNGERGQSAKVRKWTGCG---PHQODEG-RGPAGSGTRQVFSM 407
DB 86 SDRPGRCP-----APPSL-----TASGDLRVRRSGRSGPGTAPGPGXAGASP 130
QY 408 AANNKEGGT-----ASVATGDSPPSPVPLPGKPKALPGADGTFPGCPP 450
DB 131 GQRHPPGCPWPQYRKGEQVFFSPASVATGDSFVPLPGKPKALPGADGTFPGCPP 190
QY 451 GRKEKSDP-VETVMDV-VETTEAGFPQATAPQEQEIDGSKLLLMORTDVLTLGLSTR 508
DB 191 GRKEKFPDRSSWNVMGFRSNILLEAXFPQATAFP-----R 227
QY 509 LGPALKIYEH-----IKVLQGHFEDD-----DPDGFL 537
DB 228 AGNWXQIFAHAHRCASPACPSPRASPENIRAPHQGSARPLXGMWDPDGFL 280

RESULT 12
ADQ97686
ID ADQ97686 standard; protein; 801 AA.
XX ADQ97686;
XX 07-OCT-2004 (first entry)
XX Mouse cancer associated sequence MP2-10-028, SEQ ID 663.
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.
XX Mus musculus.
XX WO2004060304-A2.
XX 22-JUL-2004.
XX 22-DEC-2003; 2003WO-US041389.
XX 27-DEC-2002; 2002US-00330773.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
XX WPI; 2004-543781/52.
XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX Claim 1; SEQ ID NO 663; 199pp; English.
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 801 AA;
SQ

Query Match 11.9%; Score 340; DB 8; Length 801;
Best Local Similarity 26.6%; Pred. No. 1.9e-12;
Matches 123; Conservative 40; Mismatches 175; Indels 124; Gaps 18;

QY 4 PPALPP-PETAAAAATTAAASSSA---ASPHQEWILDTIDSLRSRKARPD-----50
DB 308 PSTPTPTPLRHAATRFATSLGSAFHPVLPHYA---TVPRLNKNSRSPSSVNTPSSQ 362
QY 51 --LERICRMVRRRHGPEPERTAELE---KLIQORAVLRVYKGSISYRNAARVQPPRR 104
DB 363 PPAAKSCAWPTSNFSLPSPIMISSPPFGKATGPRVLPVVCVSPV-----PQM 412
QY 105 GATPPAPPPA-----PRGAPAAAAAAPPPTAPPPPPAPVAAAAAPARAPAAAAATA 158
DB 413 PPSPTAPNGSLDSVTVYVSPPTSGPAAPPPPPPPPPPPPPPPPPPPPPPLPLPLASLHCSG 472
QY 159 PPSGCPAQF-----GPRQRAAPLAAP-----180
DB 473 QASPPPGTFLASTPSKPSVLFSPSAAAPASAEETPLNPELGDSSASEPCLQAASQPAESP 532
QY 181 -----PPAPAAP-----AVAPPAGPRAPPAPPAVAREPPLPPPPPPAPPQ 222
DB 533 TPQGLVLGPPAPPPPPPLPSGPAYASALPPPPG---PPPPPLPSTGPPPPPPPPPLPNQ 590
QY 223 QQQPPPPPPQPPPEGAVRAGGAA---RPVS-----LREVYVYLGSGGAGGRLT 269
DB 591 APPPPPPPPAPPLPAGSIFSGSTSEDNRPLTGLAAIAGAKLRKVRVDEGSPFGG--N 648
QY 270 RGRVQGLLEEEAARGRLERLRLGALALPRGRPCRPAPPAASARPSRKRGGE--RVLEK 328
DB 649 TGSVS--LASSKADAGR-----GNGPLPLGG--SGLMEEMSALLARRRRIAEKGTST 698
QY 329 EEEEDDEDEDEDDVSESGSEVPESDRPAGAOHQHQLNGERP 370
DB 699 EQEDRNEADAEITAKAPSTSTPEPTKRPWERTNTMNGSKSP 740

RESULT 13
AAW37151
ID AAW37151 standard; protein; 783 AA.
XX AAW37151;
XX 06-JUL-1998 (first entry)
XX Mouse neural Mena+ protein.
XX Neural Mena+ protein; mammalian Ena; Enabled protein; Evi protein;
XX cytoskeleton; cell morphology; cell adhesion; cell differentiation;
XX cell growth; cell motility; mouse.
XX Mus musculus.
XX Key Location/Qualifiers
XX Misc-difference 378 /note= "encoded by GGN"
XX FT
XX WO9801755-A1.
XX 15-JAN-1998.
XX 03-JUL-1997; 97WO-US011669.
XX 05-JUL-1996; 96US-00675815.
XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
XX Gertler FB, Soriano P, Wehland J, Niebuhr K;
XX WPI; 1998-101197/09.
XX N-PSDB; AAV02998.
XX

PT Detection of modulators of Mena and Ena-VASP-like genes and proteins -
 PT used in control of cytoskeletal dynamic events in normal and abnormal
 PS cell morphology, adhesion, motility, growth and differentiation.
 XX Example 4; Page 58-60; 77pp; English.

CC This protein comprises novel murine neural Mena+. Its amino acid sequence
 CC was deduced from a cDNA clone (see AAV02998) obtained from a mouse brain
 CC cDNA library. Neural Mena+ contains an exon that introduces 244 amino
 CC acids between amino acids 238 and 239 of mammalian Ena (Mena, see
 CC AAW37148). Two other isoforms, neural Mena++ (see AAW37152) and neural
 CC Mena+++ (see AAW37153), are also disclosed. Unlike Mena, neural Mena
 CC isoforms exhibit neural tissue-specific distribution. Based on the
 CC disclosed Mena and Evi genes (see also AAV02996-97) and proteins (see
 CC also AAW37148-49), a variety of methods and compositions are provided for
 CC screening, isolating and characterizing endogenous and exogenous factors,
 CC drugs and therapeutic agents useful to evaluate and/or control
 CC cytoskeletal dynamic events involved in normal and abnormal cell
 CC morphology, adhesion, motility, growth and/or differentiation. A method
 CC of detecting a modulator of Mena activity/expression is claimed

XX Sequence 783 AA;

Query Match 11.8%; Score 336; DB 2; Length 783;
 Best Local Similarity 26.8%; Pred. No. 3.3e-12;
 Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;

QY 4 PPALPP-PETAAATTAASSSA---ASPHYQEWILDTIDLSRSRKARPD----- 50
 DB 290 PTSTPTPLRHAATRTATSLGSAFHVLPHYA-----TVPRPLNKNRSPSPVNTPSQ 344
 QY 51 --LERICRMVRRRHGPEPERTRAELE---KLIQRAVLRSYKGSISYRNAARVQPPRR 104
 DB 345 PPAKSCAWPTSNFSLPPSPPTMISPPGKATGPRPVLPCVSSPV-----PQM 394
 QY 105 GATPPAPPRA-----PRGAPAAAAAAPPPTPAPPPAPVAAAAAPARAPRAA-----AA 154
 DB 395 PPSPTAPNGSLDSVTYVPSPPPTSGPAAPPPPPPPPPPPPPPPPPPPPLPLPLASLSHSGS 454
 QY 155 AATAPPS-----PGP-----AQPGRAPRAAPLAA 179
 DB 455 QASPPPGTPLASTPSSKPSVLPFSAGAPASAEPLNPELGDSSASEPGLQA-ASQPAES 513
 QY 180 P-----PPAAPAAPP-----AVAPPAGRRAPPVAAAREPPLPPPPPPAPP 221
 DB 514 PTQGLVLPAPPPPPPLPSGPAYASALPPPGP--PPPPPLPSTGPPPPPPPPPLN 571
 QY 222 QQQQPPPPPPPPPEGGAVRAGGAA---RPVS-----LREVVYLGSGGAGGRL 268
 DB 572 QAPPPPPPPAPPLPASGIFSGSTSDNRPLTGLAAAIAGAKLRKVSERVEDGSPFGG-- 629
 QY 269 TRGRVQGLLEEEAARGRLERLIGALALPRGDRPGRAPPAASARPSRKGEGE--RVLE 327
 DB 630 NTGSVS--LASSKADAGR-----GNGLPLGG--SGLMEEMSALLARRRIAEKGSTIE 679
 QY 328 KEEEDDDDEDEDEDDVSEGSVEPESDRPAGAHQHLNGERP 370
 DB 680 TEQEDRNEADPEITAKAPSTSTPEPTRKPKWERTNTWNGSKSP 722

RESULT 14
 AAW37152
 ID AAW37152 standard; protein; 787 AA.
 XX
 AC AAW37152;
 XX
 DT 06-JUL-1998 (first entry)
 XX
 DE Mouse neural Mena++ protein.
 XX
 KW Neural Mena++ protein; mammalian Ena; Enabled protein; Evi protein;
 KW cytoskeleton; cell morphology; cell adhesion; cell differentiation;
 KW cell growth; cell motility; mouse.

XX Mus musculus.
 OS WO9801755-A1.
 PN 15-JAN-1998.
 PD 03-JUL-1997; 97WO-US0111669.
 XX 05-JUL-1996; 96US-00675815.
 XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
 PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
 XX Gertler FB, Soriano P, Wehland J, Niebuhr K;
 XX WPI; 1998-101197/09.
 XX Detection of modulators of Mena and Ena-VASP-like genes and proteins -
 PT used in control of cytoskeletal dynamic events in normal and abnormal
 PT cell morphology, adhesion, motility, growth and differentiation.
 XX Example 4; Page 60-63; 77pp; English.

CC This protein comprises novel murine neural Mena+. Its amino acid
 CC sequence was deduced from a cDNA clone obtained from a mouse brain cDNA
 CC library. Two other isoforms, neural Mena+ (see AAW37151) and neural
 CC Mena+++ (see AAW37153), are also disclosed. Unlike mammalian Ena (Mena,
 CC see AAW37148), neural Mena isoforms exhibit neural tissue-specific
 CC distribution. Based on the disclosed Mena and Evi genes (see also
 CC AAV02996-98) and proteins (see also AAW37148-49), a variety of methods
 CC and compositions are provided for screening, isolating and characterizing
 CC endogenous and exogenous factors, drugs and therapeutic agents useful to
 CC evaluate and/or control cytoskeletal dynamic events involved in normal
 CC and abnormal cell morphology, adhesion, motility, growth and/or
 CC differentiation. A method of detecting a modulator of Mena
 CC activity/expression is claimed

XX Sequence 787 AA;

Query Match 11.8%; Score 336; DB 2; Length 787;
 Best Local Similarity 26.8%; Pred. No. 3.3e-12;
 Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;

QY 4 PPALPP-PETAAATTAASSSA---ASPHYQEWILDTIDLSRSRKARPD----- 50
 DB 294 PTSTPTPLRHAATRTATSLGSAFHVLPHYA-----TVPRPLNKNRSPSPVNTPSQ 348
 QY 51 --LERICRMVRRRHGPEPERTRAELE---KLIQRAVLRSYKGSISYRNAARVQPPRR 104
 DB 349 PPAKSCAWPTSNFSLPPSPPTMISPPGKATGPRPVLPCVSSPV-----PQM 398
 QY 105 GATPPAPPRA-----PRGAPAAAAAAPPPTPAPPPAPVAAAAAPARAPRAA-----AA 154
 DB 399 PPSPTAPNGSLDSVTYVPSPPPTSGPAAPPPPPPPPPPPPPPPPPPPPLPLPLASLSHSGS 458
 QY 155 AATAPPS-----PGP-----AQPGRAPRAAPLAA 179
 DB 459 QASPPPGTPLASTPSSKPSVLPFSAGAPASAEPLNPELGDSSASEPGLQA-ASQPAES 517
 QY 180 P-----PPAAPAAPP-----AVAPPAGRRAPPVAAAREPPLPPPPPPAPP 221
 DB 518 PTQGLVLPAPPPPPPLPSGPAYASALPPPGP--PPPPPLPSTGPPPPPPPPPLN 575
 QY 222 QQQQPPPPPPPPPEGGAVRAGGAA---RPVS-----LREVVYLGSGGAGGRL 268
 DB 576 QAPPPPPPPAPPLPASGIFSGSTSDNRPLTGLAAAIAGAKLRKVSERVEDGSPFGG-- 633
 QY 269 TRGRVQGLLEEEAARGRLERLIGALALPRGDRPGRAPPAASARPSRKGEGE--RVLE 327
 DB 634 NTGSVS--LASSKADAGR-----GNGLPLGG--SGLMEEMSALLARRRIAEKGSTIE 683
 QY 328 KEEEDDDDEDEDEDDVSEGSVEPESDRPAGAHQHLNGERP 370

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 81.6548 seconds
(without alignments)
491.841 Million cell updates/sec

Title: US-10-671-242-43
Perfect score: 2850
Sequence: 1 MAGPPALPPPTAAATAA.....HIKVLQGHFEDDDPDPGLG 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2850	100.0	538	4	US-09-616-289-43
2	2594	91.0	550	4	US-09-616-289-47
3	1445.5	50.7	317	3	US-08-979-608A-2
4	1445.5	50.7	317	4	US-09-517-849-2
5	1445.5	50.7	317	4	US-09-616-289-2
6	1191.5	41.8	252	3	US-08-979-608A-4
7	1191.5	41.8	252	4	US-09-517-849-4
8	1191.5	41.8	252	4	US-09-616-289-4
9	1170	41.1	217	3	US-08-979-608A-7
10	1170	41.1	217	4	US-09-517-849-7
11	1170	41.1	217	4	US-09-616-289-7
12	1086.5	38.1	232	3	US-08-979-608A-3
13	1086.5	38.1	232	4	US-09-517-849-3
14	1086.5	38.1	232	4	US-09-616-289-3
15	336	11.8	802	4	US-09-823-240A-2
16	319.5	11.2	783	4	US-09-252-991A-18035
17	301.5	10.6	8991	4	US-08-714-741-32
18	298.5	10.5	1958	1	US-07-945-283-2
19	298	10.5	581	4	US-09-949-016-9978
20	292.5	10.3	504	3	US-09-219-849-3
21	292.5	10.3	561	1	US-08-642-255-52
22	289.5	10.2	727	4	US-09-902-540-12383
23	288	10.1	330	1	US-08-642-255-32
24	288	10.1	408	1	US-07-609-716-65
25	288	10.1	408	3	US-08-478-029A-65
26	288	10.1	408	3	US-08-478-029A-65
27	286.5	10.1	960	3	US-09-219-849-5

28	286.5	10.1	1706	4	US-09-252-991A-31760	Sequence 31760, A
29	286.5	10.1	3122	4	US-10-237-551-201	Sequence 201, App
30	286.5	10.1	3122	4	US-10-237-551-250	Sequence 250, App
31	285.5	10.0	420	4	US-09-902-540-13993	Sequence 13993, A
32	284	10.0	357	1	US-07-609-716-66	Sequence 66, Appl
33	284	10.0	357	1	US-08-642-255-33	Sequence 33, Appl
34	284	10.0	357	3	US-08-475-411A-66	Sequence 66, Appl
35	284	10.0	357	3	US-08-478-029A-66	Sequence 66, Appl
36	282.5	9.9	720	3	US-09-219-849-4	Sequence 4, Appli
37	282.5	9.9	777	1	US-08-642-255-53	Sequence 53, Appli
38	281	9.9	905	2	US-08-574-959A-9	Sequence 9, Appli
39	281	9.9	905	3	US-09-357-014-9	Sequence 9, Appli
40	277.5	9.7	2142	4	US-09-538-092-1142	Sequence 1142, Ap
41	275.5	9.7	527	4	US-09-370-838-216	Sequence 216, App
42	275.5	9.7	527	4	US-09-854-133-216	Sequence 216, App
43	274.5	9.6	1135	2	US-08-574-959A-7	Sequence 7, Appli
44	274.5	9.6	1135	3	US-09-357-014-7	Sequence 7, Appli
45	269.5	9.5	726	4	US-09-252-991A-20675	Sequence 20675, A

ALIGNMENTS

RESULT 1
US-09-616-289-43
; Sequence 43, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-43

Query Match	100.0%;	Score	2850;	DB	4;	Length	538;
Best Local Similarity	100.0%;	Pred. No.	2.5e-172;				
Matches	538;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAGPPALPPPTAAATAA	SSSSAASHYQEWILDTIDLSR	KARPDLEICRMVR	60		
Db	1	MAGPPALPPPTAAATAA	SSSSAASHYQEWILDTIDLSR	KARPDLEICRMVR	60		
Qy	61	RHGPETRAEKLIOQRAVL	RVSKGISVRNARVQPPRRGATPPAP	PRGAPA	120		
Db	61	RHGPETRAEKLIOQRAVL	RVSKGISVRNARVQPPRRGATPPAP	PRGAPA	120		
Qy	121	AAAAAAPPPTAPPPPPA	VAAAAAATAAPSPGPAQFGPRA	QAALAAP	180		
Db	121	AAAAAAPPPTAPPPPPA	VAAAAAATAAPSPGPAQFGPRA	QAALAAP	180		
Qy	181	PPAPAPPAVAPAGPRR	APPAVAREPLPPPPQPPAPPQPP	PPPPGGAV	240		
Db	181	PPAPAPPAVAPAGPRR	APPAVAREPLPPPPQPPAPPQPP	PPPPGGAV	240		


```
QY 241 RAGGAAPVSLREVYVLGGSGAGGLTRGRVQGLLEEEAAARGRLERTRIGALALPRG 300
Db 241 RAGGAAPVSLREVYVLGGSGAGGLTRGRVQGLLEEEAAARGRLERTRIGALALPRG 300
QY 301 DRGCRAPPAASARPSRKRGGEERVLKEKEEEDDEDEDDVSGSEVPESDRPAGAQ 360
Db 301 DRGCRAPPAASARPSRKRGGEERVLKEKEEEDDEDEDDVSGSEVPESDRPAGAQ 360
QY 361 HHQNGERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFSMAANKKEGGTASVA 420
Db 361 HHQNGERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFSMAANKKEGGTASVA 420
QY 421 TGPDSFSPVPLPGKAPALPGADGTPGCGPPGRKEKPSDPVEMTMDVVEYFTAGPPEQA 480
Db 421 TGPDSFSPVPLPGKAPALPGADGTPGCGPPGRKEKPSDPVEMTMDVVEYFTAGPPEQA 480
QY 481 TAFQOEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCGHFEDDDPDGFLG 538
Db 481 TAFQOEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCGHFEDDDPDGFLG 538

RESULT 2
US-09-616-289-47
; Sequence 47, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-47

Query Match 91.0%; Score 2594; DB 4; Length 550;
Best Local Similarity 91.1%; Pred. No. 3.6e-156;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPTAAAAATT-AAAASSAASPHYQEWILDTIDSLSRSKARPDLEICRMVR 59
Db 1 MAGPPALPPPTAAAAATTAAAAASSAASPHYQEWILDTIDSLSRSKARPDLEICRMVR 60
QY 60 RRGHGPPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRARGAP 119
Db 61 RRGHGPPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRARGGP 120
QY 120 AAAAAAAPPPTAPPPPPAPV-AAAAAPARAPR-AAAAAATAPPSGPAQPGRAORAAP 176
Db 121 --AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAAATAPPSGPAQPGRAORAAP 178
QY 177 LAAPPAPAPPAVAPAGPRRAPP--AVAAREPLPPPPPPPPPPPPPPPPPPPPPPPP 229
Db 179 LAAPPAPAPPAVAPAGPRRAPPAAVAARESPLPPPPPPPPPPPPPPPPPPPPPPPP 238
```

```
QY 230 QPOPPEGAVRAGGAARPVSLREVYVLGGSGAGGLTRGRVQGLLEEEAAARGRLER 289
Db 239 QPOPPEGAAARAGGAPRVPVSLREVYVLGGSGAGGLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAPPAASARPSRKRGGEERVLKEKEEEDDEDEDD--VSEG 347
Db 299 TRIGALALPRGDRPGRAPPAASARAARNKRGAEERVLKEKEEEDDEDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN--GERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFS 406
Db 359 SEVPESDRPAGAQHQLNGERGPOTAKERAKESWLSLGGPHPGQEGRGGAAGSGTRQVFS 418
QY 407 MAAMNKEGTSASVATGPDSPSPVPLPGKAPALPGADGTPGCGPPGRKEKPSDPVEMTMD 466
Db 419 MAALSKEGGSASTTGPDSFSPVPLPGKAPALPGADGTPGCGPAGRKEKPADPVEWTMD 478
QY 467 VVEYFTAGFPQATAFQOEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCG 526
Db 479 VVEYFTAGFPQATAFQOEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCG 538
QY 527 HFEDDDPDGFLG 538
Db 539 HFEDDDPEGFLG 550

RESULT 3
US-08-979-608A-2
; Sequence 2, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-979-608A-2
```

```
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-517-849-2

Query Match      50.7%; Score 1445.5; DB 3; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-84;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 294
Db 11 PKGAARAGGAPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 70
QY 295 LALPRGDRPGARPPAASARPSRKRGGEERVLKEKEEEDDEDDED--VSEGSEVPE 352
Db 71 LALPRGDRPGARPPAASARPSRKRGGEERVLKEKEEEDDEDDED--VSEGSEVPE 130
QY 353 SDRPAGAQQHQLN--GERGPOSASAKERVKWTPCGPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQQHQLNGERGPGQAKERAKWSLCPHPGQEEGRGPAAGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGPDSPSPVPLPPGKPALPGADGTPPGCPGRKEKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTTGPDPSPSPVPLPPGKPALPGADGTPPGCPGRKEKPADPVETVMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFEED 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFEED 310
QY 532 DDPGFLG 538
Db 311 DPEGFLG 317

RESULT 4
US-09-517-849-2
; Sequence 2, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Applicant: Lees, Robert S.
; Applicant: Law, Simon W.
; Applicant: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

;
; Query Match      50.7%; Score 1445.5; DB 4; Length 317;
; Best Local Similarity 89.3%; Pred. No. 5.5e-84;
; Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;
;
QY 235 PEGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 294
Db 11 PKGAARAGGAPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 70
QY 295 LALPRGDRPGARPPAASARPSRKRGGEERVLKEKEEEDDEDDED--VSEGSEVPE 352
Db 71 LALPRGDRPGARPPAASARPSRKRGGEERVLKEKEEEDDEDDED--VSEGSEVPE 130
QY 353 SDRPAGAQQHQLN--GERGPOSASAKERVKWTPCGPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQQHQLNGERGPGQAKERAKWSLCPHPGQEEGRGPAAGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGPDSPSPVPLPPGKPALPGADGTPPGCPGRKEKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTTGPDPSPSPVPLPPGKPALPGADGTPPGCPGRKEKPADPVETVMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFEED 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFEED 310
QY 532 DDPGFLG 538
Db 311 DPEGFLG 317

RESULT 5
US-09-616-289-2
; Sequence 2, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Applicant: Lees, Robert S.
; Applicant: Law, Simon W.
; Applicant: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-616-289-2

Query Match      50.7%; Score 1445.5; DB 4; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-84;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;
;
QY 235 PEGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 294
```

Db 11 PKGAARAGAPVSLREVRVYLGSSGAGRLTRGVQGLLEEAARGLRTRLGA 70
QY 295 LALPRGDRPGRAPPAASARPSKRGGEERVLKEKEEEDDDDEDEDD--VSEGSVEPE 352
Db 71 LALPRGDRPGRAPPAASARAARNAKRAEERVLKEKEEEDDDDDVVSSEGSVEPE 130
QY 353 SDPAGAAHQLN-GERGPOSAKERVKWTFCGPHQOQDEGRGAPGSGTRQVFSMAAMN 411
Db 131 SDPAGAAHQLNGGGRGPQTAKERAKEWSLCGPHQOQDEGRGAPGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGPDSPSPVPLPPGKPPALPGADGTFPGCPGRKEKPSDPVEMTVMDVVEYF 471
Db 191 KEGGSASSTGPDSPSPVPLPPGKPPALPGADGTFPGCPGRKEKPSDPVEMTVMDVVEYF 250
QY 472 TRAGFPEQATAFQEQEIDGKSLLMQRTDVLTLGSLIRLGPALKIYEHKIKVLOQGHFEDD 531
Db 251 TRAGFPEQATAFQEQEIDGKSLLMQRTDVLTLGSLIRLGPALKIYEHKIKVLOQGHFEDD 310
QY 532 DPGFLG 538
Db 311 DPEGFLG 317

RESULT 6
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match 41.8%; Score 1191.5; DB 3; Length 252;
Best Local Similarity 88.5%; Pred. No. 4.6e-68;
Matches 223; Conservative 17; Mismatches 9; Indels 3; Gaps 2;
QY 290 TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLKEKEEEDDDDEDEDD--VSEGS 347
Db 1 TRLGALALPRGDRPGRAPPAASARAARNAKRAEERVLKEKEEEDDDDDVVSSEGS 60
QY 348 SEVPESDRPAGAAHQLN-GERGPOSAKERVKWTFCGPHQOQDEGRGAPGSGTRQVFS 406
Db 61 SEVPESDRPAGAAHQLNGGGRGPQTAKERAKEWSLCGPHQOQDEGRGAPGSGTRQVFS 120
QY 407 MAAMNKEGTTASVATGPDSPSPVPLPPGKPPALPGADGTFPGCPGRKEKPSDPVEMTVMD 466
Db 121 MAALSKEGGSASSTGPDSPSPVPLPPGKPPALPGADGTFPGCPGRKEKPSDPVEMTVMD 180
QY 467 VVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDVLTLGSLIRLGPALKIYEHKIKVLOQ 526
Db 181 VVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDVLTLGSLIRLGPALKIYEHKIKVLOQ 240
QY 527 HFEDDDDPGFLG 538
Db 241 HFEDDDDPGFLG 252
RESULT 7
US-09-517-849-4
; Sequence 4, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-517-849-4
Query Match 41.8%; Score 1191.5; DB 4; Length 252;

	Best Local Similarity	88.5%;	Pred.	No. 4.6e-68;	
	Matches	223;	Conservative	17;	Mismatches 9; Indels 3; Gaps 2;
Qy	290	TRLGALALPRGRDRFCRAPPAASARPSRSKRGGEERVLKEEEEEDEDD--VSEG	347		
Dd	1	TRLGALALPRGRDRFCRAPPAASARAARKRAGEERVLKEEEEDDEDDEDDVVSE	60		
Qy	348	SEVPESDRPAGAOHOLN-GERGPSAKERVKEWTPCGPHOQDGRGPAPSGSTRQVFS	406		
Dd	61	SEVPESDRPAGAOHQLNGGSRGPQTAKERAKEWSLCPHPQEGRGPAAAGSCTRQVFS	120		
Qy	407	MAAMNKEGGTASVATGDPSPVPLPPCKPALPGADGTFFCCPCPKRKESPDPVEWTVM	466		
Dd	121	MAALSKEGGSASTTGPDSPSVPLPPCKPALPGADGTFFCPCPKRKESPDPVEWTVM	180		
Qy	467	VVEYFTTAGPFPEQATAFOEQEIDGSKSLLMORTDLVTGLSIRLGPAIKIYEHHIKVLQQ	526		
Dd	181	VVEYFTTAGPFPEQATAFOEQEIDGSKSLLMORTDLVTGLSIRLGPAIKIYEHHIKVLQQ	240		
Qy	527	HFEDDDPDGFLG	538		
Dd	241	HFEDDDPEGFLG	252		
 RESULT 8 US-09-616-289-4 ; Sequence 4, Application US/09616289 ; Patent No. 6632923 ; GENERAL INFORMATION: ; APPLICANT: Lees, Ann M. ; APPLICANT: Lees, Robert S. ; APPLICANT: Law, Simon W. ; APPLICANT: Arjona, Anibal A. ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING ; TITLE OF INVENTION: ATHEROSCLEROSIS ; FILE REFERENCE: 10797-004001 ; CURRENT APPLICATION NUMBER: US/09/616,289 ; CURRENT FILING DATE: 2000-07-14 ; PRIOR APPLICATION NUMBER: US 09/517,849 ; PRIOR FILING DATE: 2000-03-02 ; PRIOR APPLICATION NUMBER: US 08/979,608 ; PRIOR FILING DATE: 1997-11-26 ; PRIOR APPLICATION NUMBER: US 60/031,930 ; PRIOR FILING DATE: 1996-11-27 ; PRIOR APPLICATION NUMBER: US 60/048,547 ; PRIOR FILING DATE: 1997-06-03 ; NUMBER OF SEQ ID NOS: 53 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 4 ; LENGTH: 252 ; TYPE: PRT ; ORGANISM: Oryctolagus cuniculus US-09-616-289-4					
	Query Match	41.8%;	Score	1191.5;	DB 4; Length 252;
	Best Local Similarity	88.5%;	Pred.	No. 4.6e-68;	
	Matches	223;	Conservative	17;	Mismatches 9; Indels 3; Gaps 2;
Qy	290	TRLGALALPRGRDRFCRAPPAASARPSRSKRGGEERVLKEEEEEDEDD--VSEG	347		
Dd	1	TRLGALALPRGRDRFCRAPPAASARAARKRAGEERVLKEEEEDDEDDEDDVVSE	60		
Qy	348	SEVPESDRPAGAOHOLN-GERGPSAKERVKEWTPCGPHOQDGRGPAPSGSTRQVFS	406		
Dd	61	SEVPESDRPAGAOHQLNGGSRGPQTAKERAKEWSLCPHPQEGRGPAAAGSCTRQVFS	120		
Qy	407	MAAMNKEGGTASVATGDPSPVPLPPCKPALPGADGTFFCCPCPKRKESPDPVEWTVM	466		
Dd	121	MAALSKEGGSASTTGPDSPSVPLPPCKPALPGADGTFFCPCPKRKESPDPVEWTVM	180		
Qy	467	VVEYFTTAGPFPEQATAFOEQEIDGSKSLLMORTDLVTGLSIRLGPAIKIYEHHIKVLQQ	526		

Db 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
|||||

RESULT 10

US-09-517-849-7
; Sequence 7, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match 41.1%; Score 1170; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.9e-67;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 322 EERVLEKEEEEEDEDEDDVSEGVESVPSDRPAGAHQHQLNGERGQSAKERVKEWT 381
Db 1 EERVLEKEEEEEDEDEDDVSEGVESVPSDRPAGAHQHQLNGERGQSAKERVKEWT 60
QY 382 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 441
Db 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 120
QY 442 DGTFFGCGPGRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 501
Db 121 DGTFFGCGPGRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 180
QY 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538
Db 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217

RESULT 11

US-09-616-289-7
; Sequence 7, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-7
Query Match 41.1%; Score 1170; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 8.9e-67;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 322 EERVLEKEEEEEDEDEDDVSEGVESVPSDRPAGAHQHQLNGERGQSAKERVKEWT 381
Db 1 EERVLEKEEEEEDEDEDDVSEGVESVPSDRPAGAHQHQLNGERGQSAKERVKEWT 60
QY 382 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 441
Db 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 120
QY 442 DGTFFGCGPGRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 501
Db 121 DGTFFGCGPGRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 180
QY 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538
Db 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
RESULT 12
US-08-979-608A-3
; Sequence 3, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979.608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION NUMBER: 08/979.608
; FILING DATE: 26-NOV-1997
; APPLICATION DATA:
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 27-NOV-1996
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match      38.1%; Score 1086.5; DB 3; Length 232;
Best Local Similarity 87.5%; Pred. No. 1.8e-61;
Matches 203; Conservative 17; Mismatches 5; Indels 3; Gaps 2;

QY 310 ASARPSRKGGREVLKEEEEEDEDEDD--VSEGSEVPESDRPAGAHQHLN-G 366
DB 1 ASARAARNKGAERVLKEEEEEDEDEDDDDVVSEGSEVPESDRPAGAHQHLNG 60

QY 367 ERGPQAKERVKWTCGPHQGDGREGPAGSGTRQVFSMAANKEGGTASVATGPDSP 426
DB 61 ERGPQAKERAKESLWCGPHQGEGRGPAAGSGTRQVFSMAALSKEGGSASSTTGP 120

QY 427 SPVPLPGKALPCADGTPGCGPKGKPSDPVMTVMDVYFTFTEAGPPEQATAFQEQ 486
DB 121 SPVPLPGKALPCADGTPGCGPKGKPSDPVMTVMDVYFTFTEAGPPEQATAFQEQ 180

QY 487 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 538
DB 181 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 232

RESULT 13
US-09-517-849-3
; Sequence 3, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-517-849-3

Query Match      38.1%; Score 1086.5; DB 4; Length 232;
Best Local Similarity 87.5%; Pred. No. 1.8e-61;
Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;

QY 310 ASARPSRKGGREVLKEEEEEDEDEDD--VSEGSEVPESDRPAGAHQHLN-G 366
DB 1 ASARAARNKGAERVLKEEEEEDEDEDDDDVVSEGSEVPESDRPAGAHQHLNG 60

QY 367 ERGPQAKERVKWTCGPHQGDGREGPAGSGTRQVFSMAANKEGGTASVATGPDSP 426
DB 61 ERGPQAKERAKESLWCGPHQGEGRGPAAGSGTRQVFSMAALSKEGGSASSTTGP 120

QY 427 SPVPLPGKALPCADGTPGCGPKGKPSDPVMTVMDVYFTFTEAGPPEQATAFQEQ 486
DB 121 SPVPLPGKALPCADGTPGCGPKGKPSDPVMTVMDVYFTFTEAGPPEQATAFQEQ 180

QY 487 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 538
DB 181 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 232

RESULT 14
US-09-616-289-3
; Sequence 3, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-3
```

Query Match 38.1%; Score 1086.5; DB 4; Length 232;
Best Local Similarity 87.5%; Pred. No. 1.8e-61;
Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;
QY 310 ASARPSRKGGSERVLEKEEREDDEDEED--VSEGSEVPESDRPAGAQHHOLN-G 366
DB 1 ASARAARNKRAGERVLEKEEEEEDEDDDDVVSEGSEVPESDRPAGAQHHOLNG 60
QY 367 ERGPQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFSMAANKEGGTASVATGPDSP 426
DB 61 ERGPQAKERAKESLGGPHPGQEGRGPAAGSGTRQVFSMAALSKEGGSASSTTIGPDSP 120
QY 427 SPVPLPGKALPGAGTTPGCPGKPKGSDPVETVMDVVEYFTFAGPPEQATAFQEQ 486
DB 121 SPVPLPGKALPGAGTTPGCPGKPKGSDPVETVMDVVEYFTFAGPPEQATAFQEQ 180
QY 487 EIDCKSILLMORTDVLTLGLSIRLGPALKIYEHKIKVLQQGHFEDDDPDGFLG 538
DB 181 EIDCKSILLMORTDVLTLGLSIRLGPALKIYEHKIKVLQQGHFEDDDPDGFLG 232

Search completed: September 20, 2005, 12:45:32
Job time : 82.6548 secs

RESULT 15
US-09-823-240A-2
; Sequence 2, Application US/09823240A
; Patent No. 6716597
; GENERAL INFORMATION:
; APPLICANT: Frank B. Gertler
; APPLICANT: James E. Bear
; APPLICANT: Jurgen Wehland
; APPLICANT: Joseph Loureio
; TITLE OF INVENTION: Methods and Products for Regulating Cell
; TITLE OF INVENTION: Motility
; FILE REFERENCE: M00656.70064.US
; CURRENT APPLICATION NUMBER: US/09/823,240A
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/194,564
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-823-240A-2

Query Match 11.8%; Score 336; DB 4; Length 802;
Best Local Similarity 26.8%; Pred. No. 1.4e-13;
Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;
QY 4 PPALPP-PETAAATTAASSA---ASPHYQEWILDITDLSRSRKARPD----- 50
DB 309 PTSTPTPLRHAATRFATSLGSAFHPVLPHYA-----TVPRPLNKNRSPSPVNTPPSSQ 363
QY 51 --LERICRVRRHGPBETRAELE----KLIQRAVLRSYKGSISYRNAARVQPPRR 104
DB 364 PPAKSCAMPTSNFSLPSPSPIMISSPGKATGRPVLPCVSSPV-----PQM 413
QY 105 GATPPAPPRA-----PRGAPAAAAAPPPTPAPPPAPVAAAAAPAPAPRAA-----AA 154
DB 414 PPSPTAPNGSLDSVTYVSPPTSGPAAPPPPPPPPPPPPPPPPPPPPPPLPLPLASLSHCGS 473
QY 155 AATAPPS-----PGP-----AOPGPRQRAAPLAA 179
DB 474 QASPPPTPLASTPSPSKPSVLPSPSAGAPASAEPLNPELGDSASEPGLQA-ASQPAES 532
QY 180 P-----PPAAPP-----AVAPPAGRRAPPAAVAREPPLPPPPPPAPP 221
DB 533 PTPQGLVLGPAPPPPPPLPSGPAYASALPPPPGP--PPPPPLPSTGPPPPPPPPPLEN 590
QY 222 QQQQPPPPPPQPPPEGGAVRAGAA---RPVS-----LREVRYLGGSGGAGRL 268
DB 591 QAPPPPPPPAPPPLPASGIFSGTSDNRLTLGLAAAIAGAKLRKYSRVEDGSGFPGG-- 648

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 296.987 Seconds
(without alignments)
733.538 Million cell updates/sec

Title: US-10-671-242-43

Perfect score: 2850

Sequence: 1 MAGPPALPPPTAAATAA.....HIKVLQGHFDDDPDGLG 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	2850	100.0	538	9	US-09-976-740-43	Sequence 43, Appl
2	2850	100.0	538	13	US-10-023-529-43	Sequence 43, Appl
3	2850	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
4	2850	100.0	538	15	US-10-616-187-43	Sequence 43, Appl
5	2850	100.0	538	15	US-10-671-242-43	Sequence 43, Appl
6	2594	91.0	550	9	US-09-976-740-47	Sequence 47, Appl
7	2594	91.0	550	13	US-10-023-529-47	Sequence 47, Appl
8	2594	91.0	550	13	US-10-023-523-47	Sequence 47, Appl
9	2594	91.0	550	15	US-10-616-187-47	Sequence 47, Appl
10	2594	91.0	550	15	US-10-671-242-47	Sequence 47, Appl
11	1445.5	50.7	317	9	US-09-962-055-2	Sequence 2, Appli

Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 665, App
Sequence 665, App
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 202316,
Sequence 204963,
Sequence 154970,
Sequence 130068,
Sequence 35371, A
Sequence 136002,
Sequence 2, Appli
Sequence 151853,
Sequence 140753,

ALIGNMENTS

RESULT 1

US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-43

Query Match 100.0%; Score 2850; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 4e-131;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPPALPPETAAATAAATAAASSSAASHYQEWILDTIDSLRSRKARPDLEICRMVRR 60
DB 1 MAGPPALPPETAAATAAATAAASSSAASHYQEWILDTIDSLRSRKARPDLEICRMVRR 60
QY 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRAPGAPA 120
DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRAPGAPA 120
QY 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP 180
DB 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP 180
QY 181 PPAPAPPAVAPAGPARRAPPAVAREPPLPPPOPPAPPOQOQPPPOPPPPPPGGAV 240
DB 181 PPAPAPPAVAPAGPARRAPPAVAREPPLPPPOPPAPPOQOQPPPOPPPPPPGGAV 240
QY 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300
DB 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300
QY 301 DRPGRAPPAASARPSRKGGSERVLEKEEEDDDDEDEEDDVSGSEVPESDRPAGAQ 360
DB 301 DRPGRAPPAASARPSRKGGSERVLEKEEEDDDDEDEEDDVSGSEVPESDRPAGAQ 360
QY 361 HHQNGERGQSAKERVKETPCGPHQGDGEGRPAGSGTRQVFSMAANKEGGTASVA 420
DB 361 HHQNGERGQSAKERVKETPCGPHQGDGEGRPAGSGTRQVFSMAANKEGGTASVA 420
QY 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKGKEKPSDPVETVMDVVEYFTFAGPPEQA 480
DB 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKGKEKPSDPVETVMDVVEYFTFAGPPEQA 480
QY 481 TAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOGHFEDDDPDGFLG 538
DB 481 TAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOGHFEDDDPDGFLG 538

RESULT 2
US-10-023-529-43
; Sequence 43, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-43

Query Match 100.0%; Score 2850; DB 13; Length 538;
Best Local Similarity 100.0%; Pred. No. 4e-131;

Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGPPALPPETAAATAAATAAASSSAASHYQEWILDTIDSLRSRKARPDLEICRMVRR 60
DB 1 MAGPPALPPETAAATAAATAAASSSAASHYQEWILDTIDSLRSRKARPDLEICRMVRR 60
QY 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRAPGAPA 120
DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRAPGAPA 120
QY 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP 180
DB 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP 180
QY 181 PPAPAPPAVAPAGPARRAPPAVAREPPLPPPOPPAPPOQOQPPPOPPPPPPGGAV 240
DB 181 PPAPAPPAVAPAGPARRAPPAVAREPPLPPPOPPAPPOQOQPPPOPPPPPPGGAV 240
QY 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300
DB 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300
QY 301 DRPGRAPPAASARPSRKGGSERVLEKEEEDDDDEDEEDDVSGSEVPESDRPAGAQ 360
DB 301 DRPGRAPPAASARPSRKGGSERVLEKEEEDDDDEDEEDDVSGSEVPESDRPAGAQ 360
QY 361 HHQNGERGQSAKERVKETPCGPHQGDGEGRPAGSGTRQVFSMAANKEGGTASVA 420
DB 361 HHQNGERGQSAKERVKETPCGPHQGDGEGRPAGSGTRQVFSMAANKEGGTASVA 420
QY 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKGKEKPSDPVETVMDVVEYFTFAGPPEQA 480
DB 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKGKEKPSDPVETVMDVVEYFTFAGPPEQA 480
QY 481 TAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOGHFEDDDPDGFLG 538
DB 481 TAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOGHFEDDDPDGFLG 538

RESULT 3
US-10-023-523-43
; Sequence 43, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-43

Query Match 100.0%; Score 2850; DB 13; Length 538;

[illegible]

```

RESULT 5
US-10-671-242-43
; Sequence 43, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1997-06-03
; PRIOR APPLICATION NUMBER: US 60/048,547
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-43

```

```
Query Match      100.0%; Score 2850; DB 15; Length 538;
Best Local Similarity 100.0%; Pred. No. 4e-131;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPPALPPPTAAATTAATAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
   |||||
Db 1 MAGPPALPPPTAAATTAATAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
   |||||

QY 61 RHGPEPERTRAELEKLIQRAVLRSYKGSISYRNAARVQPPRRGATPPAPRAAGAPA 120
   |||||
Db 61 RHGPEPERTRAELEKLIQRAVLRSYKGSISYRNAARVQPPRRGATPPAPRAAGAPA 120
   |||||

QY 121 AAAAAAPPPTAPPPPPAPVAAAPARAARAAAAATAPSPGPAQPGPRAQRAAPLAAP 180
   |||||
Db 121 AAAAAAPPPTAPPPPPAPVAAAPARAARAAAAATAPSPGPAQPGPRAQRAAPLAAP 180
   |||||

QY 181 PPAPAAPPAVAPAGPRRAPPPPAVAREPLPPPPQPPQPPQPPQPPQPPQPPQPPQPPQ 240
   |||||
Db 181 PPAPAAPPAVAPAGPRRAPPPPAVAREPLPPPPQPPQPPQPPQPPQPPQPPQPPQPPQ 240
   |||||

QY 241 RAGGAARPVSLRVVRYLGGSGAGGRLTRGRVQGLLEBEAAARGRLERTRLGALALPRG 300
   |||||
Db 241 RAGGAARPVSLRVVRYLGGSGAGGRLTRGRVQGLLEBEAAARGRLERTRLGALALPRG 300
   |||||

QY 301 DRPGRAPPAASARPSKSGEERVLKEEEDDEDEDEDDVSGSEVPESDRPAGAQ 360
   |||||
Db 301 DRPGRAPPAASARPSKSGEERVLKEEEDDEDEDEDDVSGSEVPESDRPAGAQ 360
   |||||

QY 361 HHQLNGERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTQVGFSAAMNKEGGTASVA 420
   |||||
Db 361 HHQLNGERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTQVGFSAAMNKEGGTASVA 420
   |||||

QY 421 TGPDSPPVPLPGKPKALPGADGTPFCGPPGKPKSDPVETVMDVVEYFTEAGPPEQA 480
   |||||
Db 421 TGPDSPPVPLPGKPKALPGADGTPFCGPPGKPKSDPVETVMDVVEYFTEAGPPEQA 480
   |||||

QY 481 TAFEQEI DGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQOGHFEDDDPDGFLG 538
   |||||
Db 481 TAFEQEI DGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQOGHFEDDDPDGFLG 538
   |||||

RESULT 6
US-09-976-740-47
; Sequence 47, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-976-740-47
```

```
Query Match      91.0%; Score 2594; DB 9; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPTAAATTAATAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 59
   |||||
Db 1 MAGPPALPPPTAAATTAATAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
   |||||

QY 60 RHGPEPERTRAELEKLIQRAVLRSYKGSISYRNAARVQPPRRGATPPAPRAAGAPA 119
   |||||
Db 61 RHGPEPERTRAELEKLIQRAVLRSYKGSISYRNAARVQPPRRGATPPAPRAAGAPA 120
   |||||

QY 120 AAAAAAPPPTAPPPPPAPV-AAAAPARAPR--AAAAATAPSPGPAQPGPRAQRAAP 176
   |||||
Db 121 --AAAAAPPPTAPPPPPAPVAAAPARAARAAAAATAPSPGPAQPGPRAQRAAP 178
   |||||

QY 177 LAAPPAPAAAPPAVAPAGPRRAPPP--AVALAREPLPPPPQPPQPPQPPQPPQPPQPP 229
   |||||
Db 179 LAAPPAPAAAPPAVAPAGPRRAPPPAAAVARESPLPPLPPPPQPPQPPQPPQPPQPPQ 238
   |||||

QY 230 QOPPPGGAAGVAGGAARPVSLRVVRYLGGSGAGGRLTRGRVQGLLEBEAAARGRLER 289
   |||||
Db 239 QOPPPGGAAGVAGGAARPVSLRVVRYLGGSGAGGRLTRGRVQGLLEBEAAARGRLER 298
   |||||

QY 290 TRLGALALPRGDRPGRAPPAASARPSKSGEERVLKEEEDDEDEDEDD--VSEG 347
   |||||
Db 299 TRLGALALPRGDRPGRAPPAASARPSKSGEERVLKEEEDDEDEDEDDVSEG 358
   |||||

QY 348 SEVPESDRPAGAQHQLN--GERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTQVFS 406
   |||||
Db 359 SEVPESDRPAGAQHQLNGERGQPTAKERAKESWLCGPHPGQEGRGPAAGSGTQVFS 418
   |||||

QY 407 MAAMNKEGGTASVATGPDSPSPVPLPGKPKALPGADGTPFCGPPGKPKSDPVETVMD 466
   |||||
Db 419 MAALSKEGGSASTTGPDSPPVPLPGKPKALPGADGTPFCGPPGKPKSDPVETVMD 478
   |||||

QY 467 VVEYFTEAGFPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQO 526
   |||||
Db 479 VVEYFTEAGFPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHHIKVLQO 538
   |||||

QY 527 HFEDDDPDGFLG 538
   |||||
Db 539 HFEDDDPDGFLG 550

RESULT 7
US-10-023-529-47
; Sequence 47, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
```

```

; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-529-47

Query Match      91.0%; Score 2594; DB 13; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPPETAAAAATT-AAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRICHMVR 59
Db 1 MAGPPALPPPPETAAAAATTAAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRICHMVR 60
QY 60 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPPRAGAP 119
Db 61 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPPRAGGP 120
QY 120 AAAAAAAPPPTAPPPPPAPV-AAAAPARAPR--AAAAATATPPSPGPAQPGPRAQAAP 176
Db 121 --AAAAAPPPTAPPPPPAPVAAAAAPARAAAAAATAAPPSPGPAQPGPRAQAAP 178
QY 177 LAAPPAPAAPAVAPAGPRRAPPP--AVAAREPPLPPPPQPPAPPPQOQO-----PPPP 229
Db 179 LAAPPAPAAPAAPAGPRRAPPPAAVAARESPPLPPPPQPPAPPPQOQOQPPPPPPQ 238
QY 230 QOQPPPEGAVRAGGAARPVSLREVRYLGGSGAGGRLTRGRVQGLLEEEAAARGRLER 289
Db 239 QOQPPPEGGAARAGGPARPVSLREVRYLGGSSGAGGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAPPAASARPSKGGSERVLEKEEEDDDDEDEEDD--VSEG 347
Db 299 TRIGALALPRGDRPGRAPPAASARAARNKRAAGEERVLEKEEEDDDDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN--GERGPOSAKERVKEWTPCGPHOGODEGRGPAKSGSTQVFS 406
Db 359 SEVPESDRPAGAQHQLNGERGPQTAKERAKESLCGPHGQEEGRGPAKSGSTQVFS 418
QY 407 MAAMKEGGTASVATGPDSPSPVPLPGKPAFGADGTFPGCPGRKEKPSDPVETVMD 466
Db 419 MAALSKEGGSASSTGTPDPSFSPVPLPGKPAFGADGTFPGCPGRKEKPADPVENTVMD 478
QY 467 VVEYFTEAGPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 526
Db 479 VVEYFTEAGPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 538
QY 527 HFEDDDPDGFLG 538
Db 539 HFEDDDPEGFLG 550

RESULT 8
US-10-023-523-47
; Sequence 47, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27

; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-523-47

Query Match      91.0%; Score 2594; DB 13; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPPETAAAAATT-AAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRICHMVR 59
Db 1 MAGPPALPPPPETAAAAATTAAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRICHMVR 60
QY 60 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPPRAGAP 119
Db 61 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPPRAGGP 120
QY 120 AAAAAAAPPPTAPPPPPAPV-AAAAPARAPR--AAAAATATPPSPGPAQPGPRAQAAP 176
Db 121 --AAAAAPPPTAPPPPPAPVAAAAAPARAAAAAATAAPPSPGPAQPGPRAQAAP 178
QY 177 LAAPPAPAAPAVAPAGPRRAPPP--AVAAREPPLPPPPQPPAPPPQOQO-----PPPP 229
Db 179 LAAPPAPAAPAAPAGPRRAPPPAAVAARESPPLPPPPQPPAPPPQOQOQPPPPPPQ 238
QY 230 QOQPPPEGAVRAGGAARPVSLREVRYLGGSGAGGRLTRGRVQGLLEEEAAARGRLER 289
Db 239 QOQPPPEGGAARAGGPARPVSLREVRYLGGSSGAGGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAPPAASARPSKGGSERVLEKEEEDDDDEDEEDD--VSEG 347
Db 299 TRIGALALPRGDRPGRAPPAASARAARNKRAAGEERVLEKEEEDDDDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN--GERGPOSAKERVKEWTPCGPHOGODEGRGPAKSGSTQVFS 406
Db 359 SEVPESDRPAGAQHQLNGERGPQTAKERAKESLCGPHGQEEGRGPAKSGSTQVFS 418
QY 407 MAAMKEGGTASVATGPDSPSPVPLPGKPAFGADGTFPGCPGRKEKPSDPVETVMD 466
Db 419 MAALSKEGGSASSTGTPDPSFSPVPLPGKPAFGADGTFPGCPGRKEKPADPVENTVMD 478
QY 467 VVEYFTEAGPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 526
Db 479 VVEYFTEAGPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 538
QY 527 HFEDDDPDGFLG 538
Db 539 HFEDDDPEGFLG 550

RESULT 9
US-10-616-187-47
; Sequence 47, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849

```

```
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-616-187-47

Query Match          91.0%; Score 2594; DB 15; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPETAAATT-AAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 59
DB 1 MAGPPALPPETAAATTAAAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
QY 60 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 119
DB 61 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 120
QY 120 AAAAAAAPTTPAPPPAPV-AAAAAPARAPR--AAAAAATAPPPSGPAQPGPRAQRAAP 176
DB 121 --AAAAAPTTPAPPPAPVAAAAAPARAPRAAAAAAATAPPPSGPAQPGPRAQRAAP 178
QY 177 LAAPPAPAAAPAVAPAGPRAPP--AFAAREPPLPPPPQPPAPPOQOQ-----PPPP 229
DB 179 LAAPPAPAAAPAVAPAGPRAPPAAAVAAARESPLPPLPPQPPAPPOQOQPPPPPPQ 238
QY 230 QOPPPPEGAVRAGGAARPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 289
DB 239 QOPPPPEGGAARAGGPAPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAAPAAASARPSKRGGEERVLEKEEEDDEDEDD--VSEG 347
DB 299 TRIGALALPRGDRPGRAAPAAASARAARKGAEERVLEKEEEDDEDEDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN-GERGQSAKERVKEWTPCGPHQODEGRGAPSGTRQVFS 406
DB 419 MAALKEGGSASTTGDSFSPVPLPPGKPALPGADGTFPGCPAGRKEKPADPVEVTMD 478
QY 467 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHKIVLQOG 526
DB 479 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHKIVLQOG 538
QY 527 HFEDDDPDGFLG 538
DB 539 HFEDDDPEGFLG 550

RESULT 10
US-10-671-242-47
; Sequence 47, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
```

```
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-671-242-47

Query Match          91.0%; Score 2594; DB 15; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPETAAATT-AAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 59
DB 1 MAGPPALPPETAAATTAAAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
QY 60 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 119
DB 61 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 120
QY 120 AAAAAAAPTTPAPPPAPV-AAAAAPARAPR--AAAAAATAPPPSGPAQPGPRAQRAAP 176
DB 121 --AAAAAPTTPAPPPAPVAAAAAPARAPRAAAAAAATAPPPSGPAQPGPRAQRAAP 178
QY 177 LAAPPAPAAAPAVAPAGPRAPP--AFAAREPPLPPPPQPPAPPOQOQ-----PPPP 229
DB 179 LAAPPAPAAAPAVAPAGPRAPPAAAVAAARESPLPPLPPQPPAPPOQOQPPPPPPQ 238
QY 230 QOPPPPEGAVRAGGAARPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 289
DB 239 QOPPPPEGGAARAGGPAPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAAPAAASARPSKRGGEERVLEKEEEDDEDEDD--VSEG 347
DB 299 TRIGALALPRGDRPGRAAPAAASARAARKGAEERVLEKEEEDDEDEDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN-GERGQSAKERVKEWTPCGPHQODEGRGAPSGTRQVFS 406
DB 419 MAALKEGGSASTTGDSFSPVPLPPGKPALPGADGTFPGCPAGRKEKPADPVEVTMD 478
QY 467 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHKIVLQOG 526
DB 479 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHKIVLQOG 538
QY 527 HFEDDDPDGFLG 538
DB 539 HFEDDDPEGFLG 550

RESULT 11
US-09-962-055-2
; Sequence 2, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
```

/ Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
/ TREATING ATHEROSCLEROSIS
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/962,055
/ FILING DATE: 24-Sep-2001
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/979,608
/ FILING DATE: 26-NOV-1997
/ APPLICATION NUMBER: US 60/031,930
/ FILING DATE: 27-NOV-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 317 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-962-055-2

Query Match 50.7%; Score 1445.5; DB 9; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 294
Db 11 PKGGAARAGGPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 70

QY 295 LALPRGDRPGRAPPAASARPSRKGGEERVLKEEEEDDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARAARNKRAAGEERVLKEEEEDDEDDDDVVSEGSEVPE 130

QY 353 SDRPAGAQQHQLN-GERGPOSAKERVKWTPCPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQQHQLNGGERGPQTAKERAKENWSLCGPHQGEGRGPAAGSGTRQVFSMAALS 190

QY 412 KEGGTASVATGPDSPSPVLPFGKPKALPGADGTPFCGPPGRKEKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTTGPDSPSPVLPFGKPKALPGADGTPFCGPPGRKEKPADPVETVMDVVEYF 250

QY 472 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 531
Db 251 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 310

QY 532 DPGDFLG 538
Db 311 DPEGFLG 317

RESULT 13
US-10-023-529-2
; Sequence 2, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS

/ Publication No. US20020194633A1
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ APPLICANT: Lees, Robert S.
/ APPLICANT: Law, Simon W.
/ APPLICANT: Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
/ TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
/ ATHEROSCLEROSIS
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/09/976,740
/ CURRENT FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: 09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
/ PRIOR FILING DATE: 1996-11-27
/ PRIOR APPLICATION NUMBER: US 60/048,547
/ PRIOR FILING DATE: 1997-06-03
/ NUMBER OF SEQ ID NOS: 53
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 317
/ TYPE: PRT
/ ORGANISM: Oryctolagus cuniculus
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(317)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-09-976-740-2

Query Match 50.7%; Score 1445.5; DB 9; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 294
Db 11 PKGGAARAGGPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 70

QY 295 LALPRGDRPGRAPPAASARPSRKGGEERVLKEEEEDDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARAARNKRAAGEERVLKEEEEDDEDDDDVVSEGSEVPE 130

QY 353 SDRPAGAQQHQLN-GERGPOSAKERVKWTPCPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQQHQLNGGERGPQTAKERAKENWSLCGPHQGEGRGPAAGSGTRQVFSMAALS 190

QY 412 KEGGTASVATGPDSPSPVLPFGKPKALPGADGTPFCGPPGRKEKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTTGPDSPSPVLPFGKPKALPGADGTPFCGPPGRKEKPADPVETVMDVVEYF 250

QY 472 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 531
Db 251 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 310

QY 532 DPGDFLG 538
Db 311 DPEGFLG 317

RESULT 13
US-10-023-529-2
; Sequence 2, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS


```
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-023-529-2

Query Match          50.7%; Score 1445.5; DB 13; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVRYLGGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGA 294
Db 11 PKGGAARAGGPAPVSLREVRYLGGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGA 70
QY 295 LALPRGDRFGRAPPAASARPSKRGGERVLEKEEEEDDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRFGRAPPAASARAARKRAGEERVELEKEEEEDDEDEDD--VSEGSEVPE 130
QY 353 SDRPAGAQHQLN--GERGPOSAKERVKWTPCGPHQGDGEGRPAGSGTRQVFSMAANV 411
Db 131 SDRPAGAQHQLNGGSRGPQTAKERAKESWLCGPHPGQEGRPAAGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGDPSPVPLPGKPKALPGADGTFPGCPGRKEKPSDPVETWMDVVEYF 471
Db 191 KEGGSASSTTGDPSPVPLPGKPKALPGADGTFPGCPGRKEKPADPVETWMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 310
QY 532 DPDGFLG 538
Db 311 DPEGFLG 317

RESULT 14
US-10-023-523-2
; Sequence 2, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02

; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-023-529-2

Query Match          50.7%; Score 1445.5; DB 13; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVRYLGGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGA 294
Db 11 PKGGAARAGGPAPVSLREVRYLGGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGA 70
QY 295 LALPRGDRFGRAPPAASARPSKRGGERVLEKEEEEDDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRFGRAPPAASARAARKRAGEERVELEKEEEEDDEDEDD--VSEGSEVPE 130
QY 353 SDRPAGAQHQLN--GERGPOSAKERVKWTPCGPHQGDGEGRPAGSGTRQVFSMAANV 411
Db 131 SDRPAGAQHQLNGGSRGPQTAKERAKESWLCGPHPGQEGRPAAGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGDPSPVPLPGKPKALPGADGTFPGCPGRKEKPSDPVETWMDVVEYF 471
Db 191 KEGGSASSTTGDPSPVPLPGKPKALPGADGTFPGCPGRKEKPADPVETWMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 310
QY 532 DPDGFLG 538
Db 311 DPEGFLG 317

RESULT 15
US-10-616-187-2
; Sequence 2, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-616-187-2

Query Match 50.7%; Score 1445.5; DB 15; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRLGGSGGAGGRLTRGRVQGLLEEEAAARGRLERTRLGA 294
Db 11 PKGARAGGPARPVSLREVVRLGGSGGAGGRLTRGRVQGLLEEEAAARGRLERTRLGA 70
QY 295 LALPRGDRPGRAPPAASARPSRKRGGEERVLKEEEEEDEDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARAARNKRAGERVLEKEEEEEDEDEDDDDVVSEGSEVPE 130
QY 353 SDRPAGAHQHLN-GERGPSAKERVKEWTPCGPHQGDGRCGAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAHQHLN-GERGPSAKERVKEWTPCGPHQGDGRCGAPGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGPDSPSPVPLPGKPALPGADGTPFCGPPGKPKSPDPVEMTMDVVEYF 471
Db 191 KEGGSASSTTGPDSPSPVPLPGKPALPGADGTPFCGPGKPKPADPVEMTMDVVEYF 250
QY 472 TEAGFPQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLOQGHFEDD 531
Db 251 TEAGFPQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLOQGHFEDD 310
QY 532 DPGFLG 538
Db 311 DPEGFLG 317

Search completed: September 20, 2005, 12:53:08
Job time : 297.987 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		Length	DB	ID	Description
		Match	*				
1	2850	100.0		538	2	Q6SPF0	Q6spfo homo sapien
2	2594	91.0		550	2	Q6SPE9	Q6spe9 oryctolagus
3	1560	54.7		285	2	Q6P1S7	Q6p1s7 homo sapien
4	1512	53.1		285	2	Q6P1R3	Q6p1r3 homo sapien
5	1269	44.5		236	2	Q6PMK4	Q6pmk4 homo sapien
6	376.5	13.2		3084	2	Q8U211	Q8uz11 pseudorabies
7	369	12.9		689	2	Q6UND9	Q6und9 rheodorae
8	364.5	12.8		676	2	Q95JC9	Q95jc9 sus scrofa
9	360	12.6		745	2	Q89X06	Q89x06 bradyrhizob
10	356.5	12.5		566	2	Q95JD1	Q95jd1 sus scrofa
11	349	12.2		3204	2	Q6X248	Q6x248 bovine herp
12	343.5	12.1		636	2	Q7SP15	Q7sp15 neurospora
13	340.5	11.9		779	2	Q6FP11	Q6fcp1 candida gla
14	338.5	11.9		511	2	Q95JD0	Q95jd0 sus scrofa
15	336	11.8		802	1	ENAH_MOUSE	Q03173 mus musculu
16	332.5	11.7		3409	2	Q6SSE6	Q6sse6 chlamydomon
17	332	11.6		3247	2	Q6S553	Q6s553 bovine herp
18	332	11.6		3247	2	Q77CD4	Q77cd4 bovine herp
19	331	11.6		659	2	Q6C7Q8	Q6c7q8 yarrowia li
20	328.5	11.5		3288	2	Q7T5D9	Q7t5d9 cercopitheci
21	324.5	11.4		664	2	Q6CDQ5	Q6cdq5 yarrowia li
22	324.5	11.4		3346	2	Q7T591	Q7t591 cercopitheci
23	323	11.3		609	2	Q9XGA4	Q9xga4 spermatozop
24	321.5	11.3		889	2	Q9F2N5	Q9f2n5 streptomyces
25	321	11.3		555	1	GP1_CHLRE	Q9fpq6 chlamydomon
26	320	11.2		518	2	Q6NTX7	Q6nxt7 brachydanio
27	320	11.2		518	2	Q7ZYX2	Q7zyx2 brachydanio
28	317	11.1		784	2	Q9CYB5	Q9cyb5 gallus galli
29	316	11.1		907	2	Q9C946	Q9c946 arabidopsis
30	314.5	11.0		675	2	Q7YUR6	Q7yur6 trypanosoma
31	314.5	11.0		1046	1	IF2_STRAW	Q82k53 streptomyces

Db	361	HHQNGERGPQSAKERVKEWTPCGPHQGDGCRGAPAGSGTRQVFSMAAMNKEGGTASVA	420
QY	421	TGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKSDPVETVMDVYVFTTEAGPPEQA	480
Db	421	TGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKSDPVETVMDVYVFTTEAGPPEQA	480
QY	481	TAFOEQEIDGKSLLMQRTDVLTLGSLRGPALKIYEHHIKVLQCGHFEDDDPDGFLG	538
Db	481	TAFOEQEIDGKSLLMQRTDVLTLGSLRGPALKIYEHHIKVLQCGHFEDDDPDGFLG	538
RESULT 2			
Q6SP69			
ID	Q6SP69	PRELIMINARY;	PRT; 550 AA.
AC	Q6SP69		
DT	05-JUL-2004	(TReMBLrel. 27, Created)	
DT	05-JUL-2004	(TReMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TReMBLrel. 27, Last annotation update)	
DE	Atherin.		
OS	Oryctolagus cuniculus (Rabbit)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX	NCBI_TaxID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Aorta;		
RA	Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;		
RL	Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY453841; AAR24088.1; -		
DR	InterPro; IPR001660; SAM.		
DR	Pfam; PF00536; SAM 1; 1.		
DR	SMART; SM00454; SAM; 1.		
DR	PROSITE; PS0105; SAM DOMAIN; 1.		
SQ	SEQUENCE 550 AA; 57019 MW; AF1CF2B780D879A7 CRC64;		
Query Match			
Best Local Similarity 91.1%; Score 2594; DB 2; Length 550;			
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;			
QY	1	MAGPPALPPETAAATTTAAAASSAASPHYQEWILDTIDSLRSKARPDLEICRMVR	59
Db	1	MAGPPALPPETAAATTTAAAASSAASPHYQEWILDTIDSLRSKARPDLEICRMVR	60
QY	60	RRHGPEPTRALEKLIQORAVLRYSYKGSISYRNAARVQPPRCATPPAPRAPRGP	119
Db	61	RRHGPEPTRALEKLIQORAVLRYSYKGSISYRNAARVQPPRCATPPAPRAPRGP	120
QY	120	AAAAAAPPPTAPPPPPAPV-AAAAAPARAPR--AAAAAATAPPSGPAQPGPRAQRAAP	176
Db	121	--AAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAAATAPPSGPAQPGPRAQRAAP	178
QY	177	LAAPPAPAAAPAVAPAGPRAPP--AVAAREPLPPPPQPPAPPOQOQ-----PPPP	229
Db	179	LAAPPAPAAAPAAAPAGPRAPPAAVAARESPLPPPPQPPAPPOQOQPPPPPPQ	238
QY	230	QOPPPGEGNVRAGGAARVSLREVYVLGGSGAGRLTRGRVQGLLEEEAAARGLER	289
Db	239	QOPPPGEGGAARAGGARVSLREVYVLGGSGAGRLTRGRVQGLLEEEAAARGLER	298
QY	290	TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLEKEEEDDDDEDEEDD--VSRG	347
Db	299	TRLGALALPRGDRPGRAPPAASARARNKXAGEERVLEKEEEDDDDDDDVVSIG	358
QY	348	SEVPESDRPAGAQHQLN--GERGPQSAKERVKEWTPCGPHQGDGCRGAPAGSGTRQVFS	406
Db	359	SEVPESDRPAGAQHQLNGERGPQAKERAKESWLCGPHGQEGRGPAAGSGTRQVFS	418
QY	407	MAAMNKEGGTASVATGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKSDPVETVMD	466
Db	419	MAALSKEGGSASTTGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKSDPVETVMD	478
QY	467	VVEYFTTEAGPPEQATFAEQEIDGKSLLMQRTDVLTLGSLRGPALKIYEHHIKVLQCG	526
Db	479	VVEYFTTEAGPPEQATFAEQEIDGKSLLMQRTDVLTLGSLRGPALKIYEHHIKVLQCG	538
QY	527	HPEDDDPDGFLG	538
Db	539	HPEDDDPDGFLG	550
RESULT 3			
Q6PIS7			
ID	Q6PIS7	PRELIMINARY;	PRT; 295 AA.
AC	Q6PIS7		
DT	05-JUL-2004	(TReMBLrel. 27, Created)	
DT	05-JUL-2004	(TReMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TReMBLrel. 27, Last annotation update)	
DE	LOC90378 protein (Fragment).		
GN	Name=LOC90378;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Wozney K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung;		
RA	Strausberg R.;		
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC030129; AAH30129.1; -		
DR	InterPro; IPR001660; SAM.		
DR	InterPro; IPR010993; SAM_homology.		
DR	Pfam; PF00536; SAM 1; 1.		
DR	SMART; SM00454; SAM; 1.		
DR	PROSITE; PS50105; SAM DOMAIN; 1.		
FT	NON_TER		
SQ	SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;		
Query Match			
Best Local Similarity 100.0%; Score 1560; DB 2; Length 295;			
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	244	GAAPVSLREVYVLGGSGAGRLTRGRVQGLLEEEAAARGRLERTRLGALALPRGDRP	303
Db	1	GAAPVSLREVYVLGGSGAGRLTRGRVQGLLEEEAAARGRLERTRLGALALPRGDRP	60
QY	304	GRAPPAASARPSKRGGEERVLEKEEEDDDDEDEDDVSEGEVPSDRPAGAQHQ	363
Db	61	GRAPPAASARPSKRGGEERVLEKEEEDDDDEDEDDVSEGEVPSDRPAGAQHQ	120
QY	364	LNGERGPOSAKERVKEWTPCGPHQGDGCRGAPAGSGTRQVFSMAAMNKEGGTASVATGP	423
Db	121	LNGERGPOSAKERVKEWTPCGPHQGDGCRGAPAGSGTRQVFSMAAMNKEGGTASVATGP	180

```

QY 424 DSPSPVPLPPGKALPGADGTPGCGPPGRKEKPSDPVWMTVMDVVEYFTFAGPPEQATAF 483
DB 181 DSPSPVPLPPGKALPGADGTPGCGPPGRKEKPSDPVWMTVMDVVEYFTFAGPPEQATAF 240
QY 484 QEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLCQGHFEDDDPDGFLG 538
DB 241 QEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLCQGHFEDDDPDGFLG 295

RESULT 4
Q6P0R3 PRELIMINARY; PRT; 285 AA.
ID Q6P0R3
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RL Strausberg R.;
DR EMBL; BC065477; AAH65477.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 285 AA; 30465 MW; PF2936CAF11F901 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7,9e-48;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 VRYLGGSGAGRLTRGRVQGLLEEEAAARGLERLRLGALALPGDRPGRAPPAASAR 313
DB 1 VRYLGGSGAGRLTRGRVQGLLEEEAAARGLERLRLGALALPGDRPGRAPPAASAR 60
QY 314 PSRSKGGGRVLEKEEEDDEDEDDVSGSVPSVPSDRPAGQHQLNGRCGQSA 373
DB 61 PSRSKGGGRVLEKEEEDDEDEDDVSGSVPSVPSDRPAGQHQLNGRCGQSA 120

```

```

QY 374 KERVKEWTPCGPHQGDGCRGAPGSGTGRQVFSMAAMNKEGGTASVATGPDSPSPVPLPP 433
DB 121 KERVKEWTPCGPHQGDGCRGAPGSGTGRQVFSMAAMNKEGGTASVATGPDSPSPVPLPP 180
QY 434 GXPALPGADGTPGCGPPGRKEKPSDPVWMTVMDVVEYFTFAGPPEQATAFQEQEIDGKSL 493
DB 181 GXPALPGADGTPGCGPPGRKEKPSDPVWMTVMDVVEYFTFAGPPEQATAFQEQEIDGKSL 240
QY 494 LLMQRTDVLTLGLSIRLGPALKIYEHKIKVLCQGHFEDDDPDGFLG 538
DB 241 LLMQRTDVLTLGLSIRLGPALKIYEHKIKVLCQGHFEDDDPDGFLG 285

RESULT 5
Q96IM4 PRELIMINARY; PRT; 236 AA.
ID Q96IM4
AC Q96IM4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE LOC90378 protein (Hypothetical protein) (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007384; AAH07384.2; -.
DR EMBL; BC080588; AAH080588.1; -.
DR HSSP; P39769; 1KW4.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match
44.5%; Score 1269; DB 2; Length 236;

```

Best Local Similarity 100.0%; Pred. No. 4e-39;		Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	303	PGRAPPAASARPSRSGRGEERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQH 362	
Db	1	PGRAPPAASARPSRSGRGEERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQH 60	
QY	363	QLNGERGPOSAKERVKEWTCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATG 422	
Db	61	QLNGERGPOSAKERVKEWTCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATG 120	
QY	423	PDSPSPVPLPPGKPPALPGADGTFPGCPGGRKEKPSDPVNTVMDVVEYFTTEAGFPPOATA 482	
Db	121	PDSPSPVPLPPGKPPALPGADGTFPGCPGGRKEKPSDPVNTVMDVVEYFTTEAGFPPOATA 180	
QY	483	FQOEIDGKSLLMQRTDVLTLGSLRGLPALKIYEHKIKVLQGHFEDDDPDGFLG 538	
Db	181	FQOEIDGKSLLMQRTDVLTLGSLRGLPALKIYEHKIKVLQGHFEDDDPDGFLG 236	
RESULT 6			
ID	Q8U211	PRELIMINARY; PRT; 3084 AA.	
AC	Q8U211;		
DT	01-MAR-2002 (T-EMBLrel. 20, Created)		
DT	01-MAR-2002 (T-EMBLrel. 20, Last sequence update)		
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)		
DE	UL36 protein.		
GN	Names=UL36;		
OS	Pseudorabies virus (strain Kaplan) (PRV).		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Alphaherpesvirinae; Varicellovirus.		
OX	NCBI_TaxID=33703;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Kaplan;		
RX	MEDLINE=21851120; PubMed=118611875;		
RX	DOI=10.1128/JVI.76.6.3065-3071.2002;		
RA	Klupp B.G., Fuchs W., Granzow H., Nixdorf R., Mettenleiter T.C.;		
RT	"Pseudorabies virus UL36 tegument protein physically interacts with the UL37 protein."		
RL	J. Virol. 76:3065-3071(2002).		
DR	EMBL; AJ422133; CAD19511.1; -		
DR	Pfam; PF04843; Herpes_tes_N; 1.		
DR	Pfam; PF03586; Herpes_UL36; 1.		
SQ	SEQUENCE 3084 AA; 324402 MW; ECAD9E1E3DC22D1A CRC64;		
Query Match 13.2%; Score 376.5; DB 2; Length 3084;			
Best Local Similarity 25.0%; Pred. No. 4.1e-06;			
Matches 159; Conservative 44; Mismatches 194; Indels 239; Gaps 29;			
QY	2	AGPPALPPPTATAAATAAASASAPHYQEWILDTIDSLRSRKARPDLER----- 53	
Db	2169	AAPPA-PPPAKPAELTPAAKLAPPAPPP-----AKPVETTLTKPQPQPPQPPAAHKK 2221	
QY	54	-----ICRMVRRHGPETRTAELEKLIQORAVLVSVKSGISYENARAVOPRR 104	
Db	2222	PAAGTKAAAPKQOPREPAPKPHSPRKIQPSLK-----ARIEPPPP 2263	
QY	105	GATPPP-----APPRAPGAPAAAAAAPPPTPAPPPPPAPVAAAAPAPAP---RA 151	
Db	2264	VINPPYPATAPETAPPEAPQAQPPNAAKTTPQGPQPPPPPPPSQAQPPAKPPAQPA 2323	
QY	152	AAAAATAPSPGPAQGPQAQRAAPLAAPPA-----PAAPPAVAPAGPRRAPPPA 203	
Db	2324	TAAATAPKATPTQPPPTTAQ--TQTAPPPPSAATAAAQVPPPPSSQAAPKRGAPP-- 2379	
QY	204	VAAREPLPPPPQ-----PPAPP-----QQQOPPPQ-----PQPPPEG 237	
Db	2380	-----APPAPPPSAQTTLPRPAAPPAPPPPSAQTTLPRPAAPPSPADATPTTPAGPAP 2435	
QY	238	GAVRAGG-----AARPSLVREWRYLGGSGGAGGLTRGRVQGLLEBEAARGLE-- 288	
Query Match 12.9%; Score 369; DB 2; Length 689;			
Best Local Similarity 31.6%; Pred. No. 2.6e-06;			
Matches 120; Conservative 23; Mismatches 133; Indels 104; Gaps 15;			
QY	97	ARVQPRRGATPPAPPRAPGAPAAAAAAPPPTTAPPPPPAPVAAAAAPARAPAAAAA 156	
Db	52	AKEAPPGRPAPAPPKAAPSAPPPPPAAAPHVAPPPPPAPPPRAAPP--PPPPAAPA 109	
QY	157	TAPPSGPAQGPRAQRAAPLAAPPAAPPAVAPPAGPRRAPPPAVAAAREPPLPPPP- 215	
Db	110	PAPKAAEPPPPPPPSHSAAPP--PPPPHAAPP--PPPAKPSAPPTAAAPERPAAPPA 163	
QY	216	-----QPPAPPQQQP-----PPQPQPP-----PPGAVRAGGAARPSLVREWRYLGGSGGA 264	
Db	164	AAPVPPAPPAPEAGAPQRRGPPFPFGAVPPNAPVPPNAA--APDAAPDAKQ-----PP 212	
QY	265	CGRLTRGRVQGLLEBEAARGRLERLRLGALALPRGDRPGR-----PPAASARP 314	
RESULT 7			
ID	Q6ND96	PRELIMINARY; PRT; 689 AA.	
AC	Q6ND96;		
DT	05-JUL-2004 (T-EMBLrel. 27, Created)		
DT	05-JUL-2004 (T-EMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (T-EMBLrel. 27, Last annotation update)		
DE	Possible OmpA family member precursor.		
GN	OrderedLocusNames=RPA0213;		
OS	Rhodopseudomonas palustris.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Bradyrhizobiaceae; Rhodopseudomonas.		
OX	NCBI_TaxID=1076;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CGA009 / ATCC BAA-98;		
RX	PubMed=14704707; DOI=10.1038/nbt923;		
RA	Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,		
RA	Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,		
RA	Gibson J.L., Hanson T.E., Bobst C., Torres Y Torres J.L., Peres C.,		
RA	Harrison F.H., Gibson J., Harwood C.S.;		
RT	"Complete genome sequence of the metabolically versatile phototrophic bacterium Rhodopseudomonas palustris.";		
RL	Nat. Biotechnol. 22:55-61(2004).		
DR	EMBL; BX572593; CAB25657.1; -		
DR	GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.		
DR	InterPro; IPR006665; OmpA/MotB.		
DR	Pfam; PF00691; OmpA; 1.		
DR	Prodom; PD000930; OmpA/MotB; 1.		
KW	Complete proteome; Signal.		
FT	SIGNAL 1 27 Potential.		
SQ	SEQUENCE 689 AA; 70745 MW; 63178BDA3E579B6B CRC64;		
Query Match 12.9%; Score 369; DB 2; Length 689;			
Best Local Similarity 31.6%; Pred. No. 2.6e-06;			
Matches 120; Conservative 23; Mismatches 133; Indels 104; Gaps 15;			
QY	97	ARVQPRRGATPPAPPRAPGAPAAAAAAPPPTTAPPPPPAPVAAAAAPARAPAAAAA 156	
Db	52	AKEAPPGRPAPAPPKAAPSAPPPPPAAAPHVAPPPPPAPPPRAAPP--PPPPAAPA 109	
QY	157	TAPPSGPAQGPRAQRAAPLAAPPAAPPAVAPPAGPRRAPPPAVAAAREPPLPPPP- 215	
Db	110	PAPKAAEPPPPPPPSHSAAPP--PPPPHAAPP--PPPAKPSAPPTAAAPERPAAPPA 163	
QY	216	-----QPPAPPQQQP-----PPQPQPP-----PPGAVRAGGAARPSLVREWRYLGGSGGA 264	
Db	164	AAPVPPAPPAPEAGAPQRRGPPFPFGAVPPNAPVPPNAA--APDAAPDAKQ-----PP 212	
QY	265	CGRLTRGRVQGLLEBEAARGRLERLRLGALALPRGDRPGR-----PPAASARP 314	


```
Db 213 GER--RGPPFG-----AFCTPNATAGMTTPPGSEAPRRGPPPPPPAAAANPPAAATP 263
QY 315 SRSRKRGGEERVLEKEBEDEDEDDVSEGEVPESE-----DRPAGAHQHLNLSRGP 370
Db 264 APS-----AAQAFTSPANPSGPAVAPVAPSGRGP 296
QY 371 QSAKERVKWTGCGPHQDGRGAPGSGTQGVFSMAAMNKEGGTASVATGPDSPSPVP 430
Db 297 QPG-----APAGGPPRPQAGPGAGAG-----PAVAPPPGQGPQVVP 333
QY 431 LPQKRALPGADGTPFCPP 450
Db 334 PQGQPPGPAVAPPAPPPP 353

RESULT 8
Q95JC9 PRELIMINARY; PRT; 676 AA.
AC Q95JC9
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RA Zhang Q., Szalay A.A., Kyeiyune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY035849; AAK61383.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005249; F:voltage-gated potassium channel activity; IEA.
DR GO: GO:0006813; P:potassium ion transport; IEA.
DR InterPro: IPR005404; KV3.3 channel.
DR PRINTS; PR01582; KV33CHANNEL.
SQ SEQUENCE 676 AA; 62297 MW; 3008BC41EFD81FC9 CRC64;

Query Match 12.8%; Score 364.5; DB 2; Length 676;
Best Local Similarity 31.7%; Pred. No. 3.7e-06;
Matches 124; Conservative 5; Mismatches 163; Indels 99; Gaps 16;

QY 101 PPRGATP--PAPP--RAPEGAP-----AAAAAAPTTPAPPPPPAPVAAAAPARA 148
Db 188 PPPPGPPPGPAPPGAPRPPGPPPPGPPPGPAPPGARPPGPPPGPPPGPAPPGARP 247
QY 149 PRAAAAAATAPSPGPAQGPRAQRAAPLAAPPAAPPAVAPVAPAG--PRRAPPVAVA- 205
Db 248 P--PGPPPLGPPPGAPPAGRPPGPPPPGPPPGPAPPGARPPGPPPPGPPPGAP 305
QY 206 --AREPPLPPPPQ--PAPPOQQOPP--PPQPPPEGGAVRAGGAAPVSLREVRYL 258
Db 306 PGARPPGPPPPGPPPGPAPPGARPPGPPPPGPPPGPAP--PPGARPPGPPPPGPP 363
QY 259 GSGGAGGRLTRGVQGLLEEEAAAGRLERLGLALPRGDRPGAPPAASARSK 318
Db 364 PGAPPGARPPPPGPP--PGPPPPGPPGPPGPPGPPGPPGPPGPPGPP 400
QY 319 RGGEERVLEKEBEDEDEDDVSEGEVPESEDRPAGAHQHLNLSRGPQSAKERVK 378
Db 401 PG-----PPPGPAPPGARPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 440
QY 379 EWTTCGPHQDGRGAPGSGTQGVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPAL 438
Db 441 PPPPAGPPPPGPPSGPAP--PGAR-----PPPGPPPPGPPPPG--PAP 480
QY 439 PGA-----DGTFFGCPGRKEKPSDP 459
Db 481 PGARPPPPGPPPPGPPPPGPPGAPPGARPPPPGPP 511
```

RESULT 9

```
Q89X06 PRELIMINARY; PRT; 745 AA.
AC Q89X06
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE B10521 protein.
GN OrderedLocusNames=blr0521;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USD110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohata M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USD110."
RL DNA Res. 9:189-197(2002).
DR EMBL: AP005936; BAC45786.1; -.
DR HSSP; P07176; 10AP.
DR GO: GO:000279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO: GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro: IPR006665; OmpA/MotB.
DR InterPro: IPR003882; Pistil_extensin.
DR Pfam: PF00691; OmpA; 1.
DR PRINTS; PR01218; PSTLEXTENSIN.
DR ProDom; PD000930; OmpA/MotB; 1.
DR Complete proteome.
SQ SEQUENCE 745 AA; 74544 MW; 155BDFCC74DBCF6D CRC64;

Query Match 12.6%; Score 360; DB 2; Length 745;
Best Local Similarity 30.7%; Pred. No. 5.7e-06;
Matches 124; Conservative 13; Mismatches 137; Indels 130; Gaps 17;

QY 100 QPPR--RGATPPAPPRAPRGAPAAA-----AAAPPTTPAPPPPPAPVAAAAPARA 148
Db 50 QPPKPGPGAAPPAAPARPAAPPPAAAPHPHPPAAAPPPAAAPPPPPPPPPPPPPPP 109
QY 149 PRAAAAAATAPSPGPAQGPRAQRAAPLAAPPAAPPAVAPVAPAGPRA-----PP 201
Db 110 PPPPPPAAPKQSPPPA---AAPQGHAPTTPPPAPPAARPAPTPPAPPAAPQHAPPPP 166
QY 202 PAVAREPPLPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 249
Db 167 PPPARPTTPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 226
QY 250 --SLREVRYLGGSGAGGRLTRGVQGLLEEEAAAGRLERLGLALPRGDRPGAP 307
Db 227 APTATPATATPATGPTGAPGAPRG-----PPPPPPPPPPPPPPPPPPPPPP 265
QY 308 -----PAASARPSRSKRGGEERVLEKEBEDEDEDEDEDEDEDEDEDEDEDEDE 360
Db 266 AAGSPAPPGATPATTTTAPG-----GTATPPSGRP----- 296
QY 361 HHQLNGERGQSAKERVKWTGCGPHQDGRGAPGSGTQGVFSMAAMNKEGGTASVA 420
Db 297 -----GPAST-----PAPGAATP---APTATPAPGALTPP 324
QY 421 TGPDSPPSPVPLPPG--KPALPGADGTPFGCAPP---GRKEKPSDP 459
Db 325 PGRPGAGTTPGQGGTTPPAGAPAGTTP-AAFPQAGGLPARPAAP 367

RESULT 10
Q95JDI
```

ID	Q95JDI	PRELIMINARY;	PRT;	566 AA.
AC	Q95JDI			
DT	01-DEC-2001	(Tremblrel. 19, Created)		
DT	01-DEC-2001	(Tremblrel. 19, Last sequence update)		
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)		
DE	Basic proline-rich protein.			
OS	Scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Parotid gland;			
RA	Zhang Q., Szalay A.A., Kyeiyune-Nyombi E., Sands J.F., Oberg K.C.,			
RA	Tieche J.-M., Leonora J.;			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY035847; AAK61381.1; -			
DR	GO; GO:0005199; F:Structural constituent of cell wall; IEA.			
DR	InterPro; IPR003882; Pistil extensin.			
DR	PRINTS; PR01218; PSTLXTENSIN.			
SQ	SEQUENCE 566 AA; 53213 MW; E33B3B5E1BDEB81A CRC64;			
Query Match 12.5%; Score 356.5; DB 2; Length 566;				
Best Local Similarity 30.5%; Pred. No. 6.2e-06;				
Matches 120; Conservative 6; Mismatches 158; Indels 109; Gaps 15;				
QY	101	PPRRGATPPAPPRAGP-----AAAAAAPPPTTAPPPPPAPVAAAPARAPAAAAAT 157		
Db	231	PPAGPPPPPPPP--PPGPPPPGAPFGARPPGPPPPGPPGPPGPPGPPGPPGPPGPPG 281		
QY	158	APSPGPAOGRQAAPLAAPAPAPAPVAPAG-----PRRAPPVAV-----AR 207		
Db	282	GPDPGPPPPGAPHGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 341		
QY	208	EPPLPPPPPP--PAPPOQOQPP-PPQOPPPGGAVRAGGAARVSLRVYLGSG 262		
Db	342	PPGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 399		
QY	263	GAGRLTRGRVQGLLBEAAAGRLRLTRGALALPRGRPGRAPPAASARPSRSKRGG 322		
Db	400	PGARPPGPPPP-----PGPPPPGAPPGARPLPGPPPPG-- 434		
QY	323	ERVLEKEEEDDEDEDDVSEGVSEVSDRPGAQHQLNGERGQSAKERVKEWTP 382		
Db	435	-----PPPGAPPGAPRPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 476		
QY	383	CGPHQOQDEGRGP-----APGSGTRQVFSMAAMNKEGGTASVATGDFSPVPLPGKP 436		
Db	477	AGPPPPPPPPGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 515		
QY	437	ALPGA-----DGTFFGCPGRKEKPSDP 459		
Db	516	APPGARPPGPPPPPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 548		
RESULT 11				
Q6X248				
ID	Q6X248	PRELIMINARY;	PRT;	3204 AA.
AC	Q6X248			
DT	05-JUL-2004	(Tremblrel. 27, Created)		
DT	05-JUL-2004	(Tremblrel. 27, Last sequence update)		
DT	05-JUL-2004	(Tremblrel. 27, Last annotation update)		
DE	UL36 very large tegument protein.			
GN	ORFNames=BHV5-20;			
OS	Bovine herpesvirus 5.			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicellovirus.			
OX	NCBI_TaxID=35244;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SV507/99;			
RX	MEDLINE=22850801; PubMed=12970418;			
RX	DOI=10.1128/JVI.77.19.10339-10347.2003;			
RA	Delhon G., Moraes M.P., Lu Z., Afonso C.L., Flores E.F., Weiblen R.,			
RA	Kutish G.F., Rock D.L.;			
RT	"Genome of bovine herpesvirus 5.";			
RL	J. Virol. 77:10339-10347(2003).			
DR	EMBL; AY261359; AAR86125.1; -			
DR	InterPro; IPR006928; Herpes teg N.			
DR	InterPro; IPR005210; Herpes_UL36			
DR	InterPro; IPR000228; RNA3' term_cycl.			
DR	Pfam; PF04843; Herpes teg N; 1.			
DR	Pfam; PF03586; Herpes UL36; 1.			
DR	PROSITE; PS01287; RTC; UNKNOWN 1.			
SQ	SEQUENCE 3204 AA; 327244 MW; 29F912A82EC57090 CRC64;			
Query Match 12.2%; Score 349; DB 2; Length 3204;				
Best Local Similarity 22.7%; Pred. No. 4.2e-05;				
Matches 158; Conservative 42; Mismatches 187; Indels 310; Gaps 26;				
QY	2	AGPPALPPETAAATTAASAAASHYQWILDTIDLSRSKARPDLERICRMVRR 61		
Db	2586	APTFVLRPE--PAPTAPASTRGAPLAAAPAEADA-----PGVER-ARLTER 2634		
QY	62	HGPEPRTAELEKLIQORAVLRVSYKSGISYRNAARVQPPRRGATPPAPPRGAPAA 121		
Db	2635	TGP-----RKSMPAALPWRRLPAAAPPAAPRTHPPPAALP---PTA 2672		
QY	122	AAAAAAPPPTPAPPPAPVAAAAAPARAPAAAAATAPSPGPAQPGPRAQAAPLAAP 181		
Db	2673	AGAAASAPSPAAPDSRTHPPPAVPPPAATAASP-PAPPPPTPS-LSPTGAPP 2730		
QY	182	PAPAAPPAVA-----PPAGPRAPPP----- 202		
Db	2731	PPPSLPPELAASTAGTGNQTVRRPGAKSMAPALPRRPPPPFPPLPSAGADGPPARG 2790		
QY	203	-----AVAREPPLP-----PPQPAPPPQQQP 226		
Db	2791	SGVDAGAAEPAPLPAVSPGPPPAVSPPPQRPSPADAGPPGAAPAPAPAPAPPAP 2850		
QY	227	PPQPQPPPEGAVAGGAARVSLRVYLGSGGAGRLTRGRVQGLLBEAAAR 286		
Db	2851	PAP-PAPPAPPSQPRADGAPP-----GPPI-----AASRN- 2880		
QY	287	LERTRLGALALPRGRPGRAPPAASARPSRSKRGGERVLEKEEEDDEDEDDVSE 346		
Db	2881	-----VPAAPAPRPTVLAIPA----- 2899		
QY	347	GSEVPESDRPA--GAQHQLNGERGQSAKERVKEWTPC-----GPHQGD----- 390		
Db	2900	----PPKPRAPKGRPGAPAKPAAGAKARRHARAVIDPTVIGSGFALSGPGADGEDADF 2955		
QY	391	-----EGRGPAGSGTRQVFSMAAMNKEGGTASVATGDPSPSPVP----- 430		
Db	2956	GVSGMYVPPDPATQAPGAPAPP-----SVAAALPAAARPAAPATQAPAGAPPSVAAA 3010		
QY	431	-----LPPGKPA-----LPGA-----DGTFFGCPGR 452		
Db	3011	PPAAARLAPPAQAPPTTRIVPAAPPPVDVGGARPTARPMARLAAPRPRESAPAAPPR 3070		
QY	453	-----KEKPSDPVETVMVVYFTEAGFPEQATAFOE 485		
Db	3071	PWASRVDLHTRDLLADAEARIRTLPAPEPPVDP-----GLDTADSSSDWSEPETEASE 3126		
QY	486	QEI-----DGKSILLMQRTDVL-----TGLSI 507		
Db	3127	TEAREGRGPRDRRARVVVANSLLIGROYLRGTGISV 3163		
RESULT 12				
Q7SF15				
ID	Q7SF15	PRELIMINARY;	PRT;	636 AA.
AC	Q7SF15			
DT	01-MAR-2004	(Tremblrel. 26, Created)		
DT	01-MAR-2004	(Tremblrel. 26, Last sequence update)		
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)		

```

DE Hypothetical protein.
GN Name=NCU07438.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrenikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysaselis M., Mauceli E., Bielke C., Rudd S., Friesman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 040-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; ABX0100025; EAA35389.1; -.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011036; PH related.
DR InterPro; IPR011960; WH1.
DR InterPro; IPR003124; WH2.
DR Pfam; PF00568; WH1; 1.
DR Pfam; PF02205; WH2; 1.
KW Hypothetical protein.
SQ
  Query Match      12.1%; Score 343.5; DB 2; Length 636;
  Best Local Similarity 30.8%; Pred. No. 2e-05;
  Matches 128; Conservative 28; Mismatches 139; Indels 121; Gaps 22;

  QY 1 MAGPALPPETAAATAAAASSAA-----SPHYQEWLDTIDSLR-----SRKA 47
  DB 268 MRAPP--PPPPAAAPRSVSESIPTSTSRGVPVPPPPARRSGKLDTENHQEPAPPFPA 325
  QY 48 RP----DLERICMVRRH-----GPEPTRALEKLIQORAVLRSYKGSISY- 93
  DB 326 VPPPIADAGKFAHSDPPRHFTSAPGPPPPRPKPTPLE-DQDPSQRFSPVPPFTGQRSVP 384
  QY 94 -----RNAARVQPPRR--GATPPAPPRAPGAPAAAAA-----APPTP--- 131
  DB 385 PPPPSRSVPPPPPNRNSAQPLLPKAPGAPPLPPASSRPPPMPLTRSPAPQAPPLP 444
  QY 132 ---APPPPPAPVAAAAAPARAAAAAATAPPSPGPAQPGRAQAAPLAAPLAPPPAPAP 188
  DB 445 TSNAPPPPLP-ATQAB---PPPLPATSPAPPPPPAPAPPPLPAAHAPPPPPPMPP 500
  QY 189 AVAPPAGPRAPPVAAAREPPLP-----PPQPPAPQOQQPPPPQPPPPGEGGAVRA 242
  DB 501 MPAPSGG---APPP-----PPPPPGMGVPPPPPPPPPPGMPPPPPAPALPPVDGSRSA 552
  QY 243 GGAARPVSLREVRYLGGSGAGS-----RLTRGVQGLLEEEAAARGRLRLTGLAL 297
  DB 553 -----VLGGITAGGIRALKKVDRSQI-----RDRSGA-AV 582
  QY 298 PRGDRPGRAP---PAASARPSRSKRGGEEVRL-----EKBEEDDDDEDEED 342
  DB 583 PGGSDSGPASSGLPPAGAAFGG--GCWADALALALQKKEKYSKSDDEDDGDWD 636

  Query Match      11.9%; Score 340.5; DB 2; Length 779;
  Best Local Similarity 25.1%; Pred. No. 3e-05;
  Matches 140; Conservative 35; Mismatches 195; Indels 187; Gaps 21;

  QY 4 PPALPPPTAAATTTAAAASSSAASPHYQEWLDTIDSLRSRKARPDLEICRVRRHG 63
  DB 156 PPSAPPPPSAPAPASPSMPSMRPKH--ESKTQNSQNSIPSAPPP-----VPGA 203
  QY 64 PEPEPTRALEKLIQORAVLRSYKGSISYRNAARVQP-----PRRGATPPAPPR---A 114
  DB 204 PPAPPTSAP-----SIPSAPAPPAPPAPPAPPMAPMAAAPPPMAAAPPA 246
  QY 115 PRGAPAAAAAAPPPTTAPPPPPAPVAAAAPA-----YLGSGSGAGGR 146
  DB 247 PPAPPAPPNNSAPSAFPAPAPPAPPMSSAPSI PALNSQSDNSPASNPSPSIPGGLPFLA 306
  QY 147 -----RAPAAAA-----AATAPPSP-GPAQPGPRA 171
  DB 307 EINARSRKGVDDKTVSHASNNASTSTNRAPPSAPPAPPPVAPVASAPSI SAPAPPI 366
  QY 172 QRA-----APLAAP-PPAPRAPP-AVAPPAGPREAPPAVAREPPLPP-PPQPPAPP 221
  DB 367 SSAPSI PSAPAPPAPPAPPAPPAPPVAAAAPPAPPAPPAPPMVSA--PPAPPAPPAPPS 424
  QY 222 QQQQPPPPQPPPEGGAVRA--GGAARPVVS-----LREVVR-----YLGSGSGAGGR 267
  DB 425 MTSAPAPAPPAPALQAASKMSKSSPPVPVSSGFLPFLAEIOKRRDDRFVVGDSNYSTK 484
  QY 268 LTRGRVQGLLEEEAAARGRLRLTGLALALPRGDRPGRAPPAPPAASARPSKRGGEEVRL 327
  DB 485 VEQTDIQG-----AAGSAPTTTSAPPPPPAGGAGMSFMSE 518
  QY 328 KEBEEDDDDEDEEDDVS-----EGSEVPESDRPAGAQQHQLNGERGPQSAKERVKEWP 382
  DB 519 IESKLKSHHHEPAPSAAPPLPPSAPOIPTSTAPP-----VP 555

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 60.5826 Seconds
(without alignments)
854.447 Million cell updates/sec

Title: US-10-671-242-43

Perfect score: 2850

Sequence: 1 MAGPPALPPPTAAATTA.....HIKVLQGHFEDDDPDGFLG 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	334.5	11.7	1585	2 T31611	hypothetical prote
2	316	11.1	907	2 E96636	hypothetical prote
3	314	11.0	464	2 S22897	extensin - volvox
4	311	10.9	505	2 S72273	actin-depolymerizi
5	298.5	10.5	1958	2 B40505	hypothetical prote
6	297.5	10.4	416	1 SKXLAG	dermal gland prote
7	297	10.4	1188	2 S49915	extensin-like prote
8	295	10.4	574	2 T43556	Wiskott-Aldrich sy
9	295	10.4	574	2 T38819	wiskott-aldrich sy
10	294	10.3	317	2 A28996	proline-rich prote
11	293.5	10.3	539	2 T28770	hypothetical prote
12	293.5	10.3	1201	2 G86441	unknown protein li
13	290.5	10.2	1733	1 B45344	probable nuclear a
14	289.5	10.2	332	1 PIHUB6	salivary proline-r
15	286.5	10.1	1621	2 T15264	hypothetical prote
16	282.5	9.9	839	2 F75518	hypothetical prote
17	279.5	9.8	1870	2 S37671	MHC class III hist
18	279.5	9.8	1872	2 S36152	MHC class III hist
19	278	9.8	660	1 Q0B33	BHLFI protein - hu
20	277.5	9.7	2142	2 B35098	MHC class III hist
21	276.5	9.7	416	2 T34279	hypothetical prote
22	276	9.7	2715	2 T13049	eyelid - fruit fly
23	274	9.6	3149	1 Q0B38	BPLFI protein - hu
24	273.5	9.6	439	2 S51939	chitinase (EC 3.2.
25	271	9.5	708	2 D96711	hypothetical prote
26	270.5	9.5	801	2 T29018	hypothetical prote
27	269.5	9.5	240	2 B24264	proline-rich prote
28	268.5	9.4	1151	2 T18535	high molecular mas
29	268.5	9.4	3534	2 T42567	tegument protein 2

30	264	9.3	589	2 T29299	hypothetical prote
31	263	9.2	760	2 T06291	extensin homolog T
32	262.5	9.2	461	2 T10741	extensin-like prot
33	262	9.2	1460	1 EDBE1F	immediate-early pr
34	261.5	9.2	775	1 EDBE11	immediate-early pr
35	261.5	9.2	969	2 T15446	hypothetical prote
36	260.5	9.1	5262	2 T03454	ALR protein - huma
37	259.5	9.1	309	2 S10889	proline-rich prote
38	259.5	9.1	310	1 PIHUSD	salivary proline-r
39	259.5	9.1	1100	2 JC8033	leukocyte formin p
40	259	9.1	666	2 B70803	hypothetical prote
41	258	9.1	496	2 T17908	proline/lysine-ric
42	257	9.0	599	2 T10798	phosphorin-S - Vo
43	257	9.0	707	2 A46302	PTB-associated spl
44	255.5	9.0	1611	2 T38236	hypothetical prote
45	255.5	9.0	1638	2 A42091	transcription acti

ALIGNMENTS

RESULT 1

T31611

hypothetical protein Y50E8A.g - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31611

R:Steward, C.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z21047

A:Accession: T31611

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1585 <WIL>

A:Cross-references: EMBL:AL117200; MID:e1549770; PIDN:CAB55050.1; CESP:Y50E8A.g

A:Experimental source: clone Y50E8A

C:Genetics:

A:Gene: CESP:Y50E8A.g

A:Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match					11.7%; Score 334.5; DB 2; Length 1585;
Best Local Similarity					26.3%; Pred. No. 2.7e-07;
Matches 124; Conservative 17; Mismatches 155; Indels 175; Gaps 14;					
QY	89	GSISYRNAARVOPRRGATPPAPPRAPRCAPA	-----A 121		
Db	907	GGVSGSALPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
QY	122	AAAAAAPPPTPPPPPPPPVAAAAPARAPRA	-----AAAAATAPPSPGPAQPPRAQRAAP	176	
Db	967	GSAAPPPPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
QY	177	LAAPPAPAP -PAVAPP	-----AGPRAPPVAAAREPPL 211		
Db	1027	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
QY	212	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
Db	1081	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
QY	236	-----EGGAVRAGGAARPVSLREVLYLGGSGGAGGRLTRGRVQGLLEBEAARGLERT	290		
Db	1141	APAPSSGG--YSGGSS	-----CGSAAGGGGGSGGYSG	1171	
QY	291	RIGALALPGDRPPGRAPPAASARPSKFGBEERVLKEEEDDEDEDEDDVSGSEV	350		
Db	1172	--GSAAPPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
QY	351	PESDRPAGAHQLNGERCPSAKERVKWTTCGPHQGDGEGRPAAGSGTRQVFSMAAM	410		
Db	1206	-----GSAAGGGGGSGGYSGSAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP		
QY	411	NKEGG-----TASVATGDFSPVPLPPGKPA---	LPGADTTPGCPFG 451		

Db 1260 AAGGGGSSGGYTGGAAPPPPPPPPPPPPPAPAPAPSSGGYSGSSGG 1310

RESULT 2

E96636

hypothetical protein T7P1.21 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E96636

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96636

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-907 <STO>

A:Cross-references: UNIPROT:Q9C946; GB:AE005173; NID:g6751696; PIDN:AAF27679.1; GSPDB:GN

C:Genetics:

A:Gene: T7P1.21

A:Map position: 1

Query Match 11.1%; Score 316; DB 2; Length 907;

Best Local Similarity 29.4%; Pred.No. 1.1e-06;

Matches 112; Conservative 39; Mismatches 130; Indels 100; Gaps 19;

QY 33 EWILDTIDSLRSRKARPDLERICRMVRRRHGPPERTRAELEKLI---QORAVLRVSYKG 89

Db 343 EDVISVVEIKQKK--DEIESIDVKMETE-----ESVNLDEESVVLNGEQDTIMKISSLE 395

QY 90 SIS-----YRNAARVQPPRRGATPPAP-----PRPRGAPAA-AAAAAPP 128

Db 396 STSESKLNHSEKVENSQLFPFPPPPPPPLSFXTASLPLSPPTPTPIADIALSMPP 455

QY 129 PTAPPPPPA--PVAAAAAPARAAAAAT----APPSPGPAQCPRAQRAAPLAAPP 182

Db 456 PPPPPPPPAVMLKHFAPPPPPPLPPAVMLKHFAPPPPTTFAFKLKGSAAPPPPPPP 515

QY 183 AP---AAPP-----AVAPPAGRRAPPVAAAREPLPP-----PPOPPAPPOQ 224

Db 516 LPTTIAAPPPPPPPRAAVAPPPPP--PPPGTAAPPPPPPPPGTCAAPPPPPPPMQN 572

QY 225 Q-----PPPPQP-----QPPEGGAVRAGGAAPVSLREVVRVYLG 259

Db 573 RAPSPPPMPMNGSGSGPPPPPPMPMLANGATPPPPPPPMAMANGAAGPPP--PPPRMGW 630

QY 260 GSGGAGRLTRGRVQGLLEEEAAARGLER--TLGLAL--ALPRGDRPGRAPPA----- 309

Db 631 ANGAAGPPPPPPGAARSLRPKKAAT--KLKRSITOLGNLYRLKKGVEGRDPNKTGSGSR 688

QY 310 ---ASAPRSKRGGSEERVLE 327

Db 689 KAGAGSAPAGKGGMADALAE 709

RESULT 3

S22697

extensin - Volvox carteri (fragment)

C:Species: Volvox carteri

C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004

C:Accession: S22697; S21006

R:Ertl, H.; Hallmann, A.; Wenzl, S.; Sumper, M.

EMBO J. 11, 2055-2062, 1992

A:Title: A novel extensin that may organize extracellular matrix biogenesis in Volvox ca

A:Reference number: S22697; MUID:92289669; PMID:1600938

A:Accession: S22697

A:Molecule type: mRNA

A:Residues: 1-464 <HAL>

A:Cross-references: UNIPROT:Q41645; EMBL:X65165; NID:g21991; PIDN:CAA46283.1; PID:g21992

C:Keywords: glycoprotein

Query Match 11.0%; Score 314; DB 2; Length 464;

Best Local Similarity 31.1%; Pred.No. 7.6e-07;

Matches 82; Conservative 18; Mismatches 86; Indels 76; Gaps 10;

QY 5 PALPPPETAAATTAATAASSSAASHYQEWILDTIDSLRSRKARPDLERICMVRRRHGP 64

Db 220 PSPPPPRVSTSPPPARVSSSPPP-----ATRSPPP-----RRITSP 257

QY 65 EPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPAPRGAPAAAAA 124

Db 258 SPVLT-----ASPLPKTSPPPPPPPPPPPPVAS 288

QY 125 AAPPPTP---APPPPPAPVAAAAAPARAAAA---AATAPPSPGPAQCPRAQRAAP 176

Db 289 PPPPPPPRVSPSPPPPPQPVVSSPPPPPPRPSPPPPPSPPSPPPPPPPPPRPS 348

QY 177 LAAPP---PAPAAPPAVAPPAGRRAPPVAAAREPLPPPPPPQPPAPPOQOQPP----- 227

Db 349 SPPPPSSESP--PPPVVSPPPPPPRASPPPPPPASPPPPPPPPPPPPPPPPATAAAN 407

QY 228 PPQOPPPPEGGAVRAGG---AARP 248

Db 408 PPSAP-----SRSRAGGPPPLGTRP 427

RESULT 4

S72273

actin-depolymerizing protein N-WASP, brain - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 23-Apr-1998 #sequence_revision 01-May-1998 #text_change 09-Jul-2004

C:Accession: S72273

R:Waki, H.; Miura, K.; Takenawa, T.

EMBO J. 15, 5326-5335, 1996

A:Title: N-WASP, a novel actin-depolymerizing protein, regulates the cortical cytoskeleton

A:Reference number: S72273; MUID:97050838; PMID:8895577

A:Accession: S72273

A:Molecule type: mRNA

A:Residues: 1-505 <MIK>

A:Cross-references: UNIPROT:Q95107; EMBL:D67066; NID:g1644231; PIDN:BAAL1082.1; PID:d101

A:Experimental source: brain

Query Match 10.9%; Score 311; DB 2; Length 505;

Best Local Similarity 31.4%; Pred.No. 1.1e-06;

Matches 93; Conservative 32; Mismatches 107; Indels 64; Gaps 13;

QY 74 EKLIQORAVLRVSYK-----GSIYRNAARVQ-----PPRRGATPPAPPAPRGAPAAA 122

Db 243 EAQLKDRETSKVIYDFIEKTGGVEAVKVELRQAPPPPPSRGGPPPPPP----- 292

QY 123 AAAAPPTTAPPPPPAPVAAAAAPARAAAAATATPPSPGPAQCPRAQRAAP--LAAPP 181

Db 293 ----PPHSSGPPPPPARGARGAPP---PPPSRAPTAAPPPPPSRPGVGAAPPPPPNRMYP 345

QY 182 PAPAAPPVAPPAGRRAPPVAAAREPLPPPPPPQPPAPPOQOQPPPPPPPEGAVR 241

Db 346 PLPALPSSA--ESGPPPPPPPLSVSGSVAPPPPPPPPPPPP-----GPPPPGGLPSGDHGV 399

QY 242 AGGAARPVSLRWRYLGGSGGAGRLTRGRVQGLLEEA---AAGR---LERTRLGAL 295

Db 400 PTPAGSKAALLDQIR-----EGALKK-----VEQNSRPVSCSRDALLDQIRQGIQ 446

QY 296 ALPRGDRPGRAPPAASARPSRKRGSEERVLBK-----EEEEDEDEDEED 343

Db 447 LKSVTDAPSTTPAPA--PTSGIVGALMEVMQKRSKAIIHSDDEDEDEDEDED 500

RESULT 5

B4050S
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhouser or Becker)
C:Species: suid herpesvirus 1
C>Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C:Accession: B4050S
R;Cheung, A.K.
J. Virol. 65, 5260-5271, 1991
A>Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A:Reference number: A40505; MUID:91374576; PMID:1654441
A:Accession: B4050S
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1958 <HE>
A:Cross-references: UNIPROT:Q69340; GB:M57505; NID:g334066; PIDN:AAA47468.1; PID:g334068
C:Superfamily: pseudorabies virus 1 nuclear antigen

	Query Match	10.5%	Score 298.5;	DB 2;	Length 1958;
	Best Local Similarity	25.1%;	Pred. No. le-05;	Mismatches 31;	Gaps 24;
	Matches 124;	Conservative			
Qy	61 RHGPEPRTAELEKLIQQRAVLVSVYSGISITYNAAARVQPRRGATPPAPPAPRCAPA	120	:	:	:
Dd	262 RQGPAIDAVQT-----LVHARQLRLGR-DVOADGA VHDGPSSN-SHGFP-GPSGPT	315	:	:	:
Qy	121 AAAAAPTTPA-----PPPFAVVAA-----APARA-----PRAAAA	154	:	:	:
Dd	316 SHODPPPGPTPTSTSHHHHHHQGPPTS PRPSTSSHODPPGGPPTSAETHHHHQDPGCGP	375	:	:	:
Qy	155 AATA-----PPSPGAQPGR-----AQRAAPLAAPPAPPAVPAPVAGPR----	197	:	:	:
Dd	376 PSTSSHHHQDPPGCGPPPPPREPSTSSSSHQGPSTRPPPPQRPPRPWPPSPQKISE	435	:	:	:
Qy	198 -----RAPPFAVAARSPPPLP	212	:	:	:
Dd	436 TRAGESTAQTLFSHSNKULFHPMGEGGEDGTAGCEDDDDRPPSP-----PPRP	499	:	:	:
Qy	213 PPPOPPAPPOOOPPPPOPPPPPEGGAVRAGGAARPVSREVRYLGGSGGAGRILTRGR	272	:	:	:
Dd	490 PPLLP PPPP-----RRRGSGGGPPGR--GGR	529	:	:	:
Qy	273 VOGLLEEAAARGKLTRLICALALPGDRPGRAPPAASARPRS KRKGSEERVLEKEEEE	332	:	:	:
Dd	530 RRG-----GKR-----RRAEGETAAAAADAEEEEE	552	:	:	:
Qy	333 DDEDEDDEEDDSVGSB-VPESDRPAG-----AQHLQNGERGQS AKER-----	376	:	:	:
Dd	553 DGDEDEDERAEDREGDGEGGPRGAGGAGESESSESRAEGAPRSAEQGVGVAGVTLG	612	:	:	:
Qy	377 ---VKEMTTCGPH-QGDQEGRGAPGSGTRQVFMSMAANKNE---GGTA SVAYGPDSPSP	428	:	:	:
Dd	613 LLVVD---GLHLDDGERAAGPA-----VAAEAADDLRGEVLPVLAP-----	653	:	:	:
Qy	429 VPLPPGPKPALPGADG	443	:	:	:
Dd	654 -PGARGPVGLHGAAG	667	:	:	:

RESULT 6

SXLLAG
dermal gland protein APeg precursor - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S07458; A34140
R;Gmachl, M.; Berger, H.; Thallhammer, J.; Kreil, G.
FEBS Lett. 260, 145-148, 1990
A>Title: Dermal glands of Xenopus laevis contain a polypeptide with a highly repetitive
A:Reference number: A34140; MUID:90127399; PMID:2298293
A:Accession: S07458
A:Molecule type: mRNA
A:Residues: 1-416 <GNA>
A:Cross-references: UNIPROT:P17437; EMBL:X51394; NID:g64547; PIDN:CAA35759.1; PID:g13343
A>Note: the authors translated the codon TTC for residue 8 as Pro

Db 719 PPPLSKTPAPPPPLSKTPVPPPPGLGRGTS SGPPP-----LGAKGS----- 761

Qy 273 VQGLLSEEAARGLRERTRIGALALPRG---DRPGRAPPAASARP-----SRSKRGGBE 323

Db 762 -NAPPPPPPPAGRGK-----ASLGLGRGVSVPAAKKTALKPLHWSKVTRAAKGS-- 812

Qy 324 RVLEKEEEEEDEDEEDVSEGSEV--PESDRPAGAHQHLNGERGQPSAK-ERVK 378

Db 813 --LWADTQKQENQPRAPEDISELSFSAVSDTTA-----KKSTGRRGSSISKPEKVQ 864

RESULT 13

B45344

probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)

C;Species: suid herpesvirus 1

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: B45344

R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.

Virology 179, 365-377, 1990

A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented ORF

A;Reference number: A45344; MUID:91021039; PMID:2171211

A;Accession: B45344

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-1733 <VLC>

A;Cross-references: UNIPROT:P33485; GB:M34651; NID:G334070; PIDN:AAA47471.1; PID:G334072

C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 10.2%; Score 290.5; DB 1; Length 1733;

Best Local Similarity 23.7%; Pred. No. 2.1e-05;

Matches 112; Conservative 27; Mismatches 121; Indels 213; Gaps 18;

Qy 108 PPAPPRAPRG--APAAAAAAPPPTAPP-----PPAPVAAA----- 143

Db 70 PPHPPRPQDHRFTEARDHRDPRDLPTTRTRDHQHRPPPTTTTITKDQHPQDPLL 129

Qy 144 -----APAPRAA-----AAAATA----- 170

Db 130 LPTKTLQEDPHLLRPTRDPPSAKTHHHQDPGPGPSTSSHHHQDPGPGPPPPRP 189

Qy 171 -----AQRAAPLAAPPAPPAAPVAPPA----- 194

Db 190 STSSSSHQGPSTRTPPPPQRPWPPPSPQKISETRAGSENTAQLFSHSENKLFSPH 249

Qy 195 -----GPRAPPAPVAAREPPLPPPPQPPAPQOQQPPPPQPPPPPEG 238

Db 250 MGGEGEGDRGTAGGEGDRDPPF---PSPPPRPPLPPPP-----PPPPPPQPPPAG 301

Qy 239 AVRAGGAARPVSLREVVRVYLGSGGAGGLTRGVRVQGLLEEAARGLRERTRIGALALP 298

Db 302 SARRR-----RRGGPPGRRRGK----- 323

Qy 299 RGRPGRAPPAASARPSRKRGEERVLEKEEEDDEDEDDVSEG-----SEVPESDR 355

Db 324 -----RRRAEGTEAAAADAEEDGDGDEDEDAEGEGREDGEGRP 367

Qy 356 PAGAQHQLNGERG-----PQSAKER-----VKWETPCGPH-QQDQGRGPAPG 398

Db 368 GAGGAGESSESGRAEGAPRSAEQVGVAGVLGLLVVRD-----GLHLDGPFERAAAGPA-- 421

Qy 399 SGTQVFSMAAMNKE-----GGTASVATGDPSPVPLPGKALPGADG 443

Db 422 -----VAAADADLLHVPVLVLAGAGPGARGPRGVLGHGAPGGADAGLEG 466

RESULT 14

PIHUB6

salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - human

N;Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H

C;Species: Homo sapiens (man)

C;Date: 04-Dec-1986 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004

C;Accession: B40750; A40750; A40750; C25372; S02127; S02127; A03293; A90502; A91974; A09

R;Azen, E.A.; Latreille, P.; Niece, R.L.

Am. J. Hum. Genet. 53, 264-278, 1993

A;Title: PRB1 gene variants coding for length and null polymorphisms among human salivary

A;Reference number: A40750; MUID:93304421; PMID:8317492

A;Accession: B40750

A;Molecule type: DNA

A;Residues: 35-392 <AZE>

A;Cross-references: UNIPROT:P04280; UNIPROT:P04281; UNIPROT:P02811; UNIPROT:Q16038; GB:S6

A;Experimental source: subject C.J. (large allele)

A;Accession: C40750

A;Molecule type: DNA

A;Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-392 <I

A;Cross-references: GB:S62929

A;Experimental source: subject M.V.O. (large allele)

A;Accession: A40750

A;Molecule type: DNA

A;Residues: 35-183,245-270,'Q',272-392 <AZ3>

A;Cross-references: GB:S62928

A;Experimental source: subject C.J. (medium allele)

A;Note: authors translated the codon CAA for residue 272 as Arg

R;Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.

J. Biol. Chem. 260, 11123-11130, 1985

A;Title: Differential RNA splicing and post-translational cleavages in the human salivary

A;Reference number: A92492; MUID:85289325; PMID:2993301

A;Accession: C25372

A;Molecule type: mRNA

A;Residues: 1-183,245-392 <MAE>

A;Cross-references: GB:K03204; NID:G190485; PIDN:AAA60185.1; PID:G190486

A;Note: alternatively splice forms lacking portions of the repeat region were also found

R;Lyons, K.M.; Stein, J.H.; Smithies, O.

Genetics 120, 267-278, 1988

A;Title: Length polymorphisms in human proline-rich protein genes generated by intragenic

A;Reference number: S02127; MUID:89121440; PMID:2851479

A;Accession: S02128

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>

A;Cross-references: EMBL:X07517

A;Accession: S02127

A;Status: translation not shown

A;Molecule type: DNA

A;Cross-references: EMBL:X07516

R;Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P.

Biochemistry 25, 2387-2392, 1986

A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent structu

A;Reference number: A90502; MUID:86243355; PMID:3521730

A;Accession: A03293

A;Molecule type: protein

A;Residues: 17-38,'P',41-51,92-148,'R',150-152 <KA2>

A;Note: among nine basic proline-rich peptides isolated from the saliva, this peptide is

A;Accession: A90502

A;Molecule type: protein

A;Residues: 275-336,'S',338-392 <KAU>

R;Saitoh, E.; Isemura, S.; Sanada, K.

J. Biochem. 94, 1991-1999, 1983

A;Title: Further fractionation of basic proline-rich peptides from human parotid saliva

A;Reference number: A91974; MUID:84161824; PMID:6671974

A;Contents: P-H

A;Accession: A91974

A;Molecule type: protein

A;Residues: 'S',338-392 <SAI>

R;Azen, E.; Lyons, K.M.; McGonigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Vanin,

Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984

A;Reference number: A94005; MUID:84298176; PMID:6089212

A;Accession: A05261

A;Molecule type: DNA

A;Residues: 35-39,'P',41-84,'G',86,'R',87-154,'R',218-246;300-306,'T',308-329,'C',331-384

A;Accession: A05262

A;Molecule type: DNA

A;Residues: 'N',57-59,'A',61-69;334-336,'S',338-339,'R',341-392 <AZ5>

R;Kauffman, D.; Wong, R.; Bennick, A.; Keller, P.

Biochemistry 21, 6558-6562, 1982

A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent structu

Job time : 63.5826 sec

THIS PAGE BLANK (USP 10)

CC This polypeptide comprises novel human low density lipoprotein (LDL)
 CC binding protein LBP-2 that is capable of binding both native and methyl
 CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
 CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human
 CC LBPs (see AAV49037-42) are claimed. An abnormality in an aspect of LBP
 CC metabolism or structure is diagnostic of a risk for atherosclerosis. The
 CC invention provides: methods for determining if an animal is at risk for
 CC atherosclerosis (e.g. for prenatal screening); methods for treating
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
 CC methods for treating a cell having an abnormality in LBP structure or
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
 CC well as recombinant vectors and host cells used to produce recombinant
 CC LBP
 XX
 SQ Sequence 217 AA;

Query Match 100.0%; Score 1170; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 4.6e-93;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EERVLEKEEEDDDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 60
 DB 1 EERVLEKEEEDDDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 60
 QY 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 120
 DB 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 120
 QY 121 DGTFFGCPGRKEKPSDPVETWMDVVEYFTAGFPEQATAFQEQIDGKSLLLMORTDV 180
 DB 121 DGTFFGCPGRKEKPSDPVETWMDVVEYFTAGFPEQATAFQEQIDGKSLLLMORTDV 180
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
 DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 2
 AAB82803
 ID AAB82803 standard; protein; 217 AA.
 XX
 AC AAB82803;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Human low density lipoprotein binding protein 2 (LBP-2).
 XX
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US0006356.
 XX
 PR 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 DR WPI; 2001-565505/63.
 DR N-PSDB; AAV326494.
 XX
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX
 XX Claim 13(g); Fig 7B; 143pp; English.

XX
 CC The present sequence is that of the N-terminal portion of novel human low
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
 CC deduced from an isolated partial cDNA clone (see AAV26494). A full-length
 CC sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP
 CC polypeptides of the invention that are capable of binding to native and
 CC methylated low density lipoproteins. Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs. Methods
 CC of determining if an animal is at risk for atherosclerosis, and methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX
 SQ Sequence 217 AA;

Query Match 100.0%; Score 1170; DB 4; Length 217;
 Best Local Similarity 100.0%; Pred. No. 4.6e-93;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EERVLEKEEEDDDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 60
 DB 1 EERVLEKEEEDDDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 60
 QY 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 120
 DB 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPALPGA 120
 QY 121 DGTFFGCPGRKEKPSDPVETWMDVVEYFTAGFPEQATAFQEQIDGKSLLLMORTDV 180
 DB 121 DGTFFGCPGRKEKPSDPVETWMDVVEYFTAGFPEQATAFQEQIDGKSLLLMORTDV 180
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
 DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 3
 AAB82806
 ID AAB82806 standard; protein; 538 AA.
 XX
 AC AAB82806;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Human low density lipoprotein binding protein 2 (LBP-2).
 XX
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US0006356.
 XX
 PR 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 DR WPI; 2001-565505/63.
 DR N-PSDB; AAV26499.
 XX
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX
 XX Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein
 CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the
 CC coding region of isolated genomic DNA (see AA26499). It differs from the
 CC sequence predicted from an LBP-2 cDNA clone (see AA82803) by the
 CC presence of an additional 321 amino acids at the N-terminus (the cDNA
 CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP
 CC polypeptides of the invention that are capable of binding to native and
 CC methylated low density lipoproteins. Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs.
 CC Polypeptides having amino acid residues 329-343, 329-354, or 529
 CC -538 (see AA82809-12) of the present sequence are claimed. Methods of
 CC determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX Sequence 538 AA;
 SQ
 Query Match 100.0%; Score 1170; DB 4; Length 538;
 Best Local Similarity 100.0%; Pred. No. 1.4e-92;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EERVLEKEEEDDEDEDDVSEGVESDPDRPAGAHQHLNGERGPOSAKERVKEWT 60
 DB 322 EERVLEKEEEDDEDEDDVSEGVESDPDRPAGAHQHLNGERGPOSAKERVKEWT 381
 QY 61 PCGPHQGDGRGPAFGSGTRQVFSMAAMNKEGTSATGPDSPSPVPLPPGKALPGA 120
 DB 382 PCGPHQGDGRGPAFGSGTRQVFSMAAMNKEGTSATGPDSPSPVPLPPGKALPGA 441
 QY 121 DGTFFGCPGRKEKPSDPVETWMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQRTDV 180
 DB 442 DGTFFGCPGRKEKPSDPVETWMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQRTDV 501
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFEDEDDPDGFLG 217
 DB 502 LTGLSIRLGPALKIYEHHIKVLQOQHFEDEDDPDGFLG 538
 RESULT 4
 AAB58957
 ID AAB58957 standard; protein; 241 AA.
 AC AAB58957;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.
 XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antitumor; antitumor; antitumor;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005881.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Rosen CA, Ruben SM;
 XX WPI; 2000-611515/58.
 DR N-PSDB; AAF21860.
 DR
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 XX neurological diseases.
 PS Claim 11; Page 1112; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 CC antiinflammatory; antitumor; antitumor; antitumor; antitumor;
 CC antifungal; antiparasitic and cardiac activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemia; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases
 XX Sequence 241 AA;
 SQ
 Query Match 99.2%; Score 1161; DB 3; Length 241;
 Best Local Similarity 99.1%; Pred. No. 3.1e-92;
 Matches 215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EERVLEKEEEDDEDEDDVSEGVESDPDRPAGAHQHLNGERGPOSAKERVKEWT 60
 DB 25 EERVLEKEEEDDEDEDDVSEGVESDPDRPAGAHQHLNGERGPOSAKERVKEWT 84
 QY 61 PCGPHQGDGRGPAFGSGTRQVFSMAAMNKEGTSATGPDSPSPVPLPPGKALPGA 120
 DB 85 PCGPHQGDGRGPAFGSGTRQVFSMAAMNKEGTSATGPDSPSPVPLPPGKALPGA 144
 QY 121 DGTFFGCPGRKEKPSDPVETWMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQRTDV 180
 DB 145 DGTFFGCPGRKEKPSDPVETWMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQRTDV 204
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFEDEDDPDGFLG 217
 DB 205 LTGLSIRLGPALKIYEHHIKVLQOQHFEDEDDPDGFLG 241
 RESULT 5
 AAB82799
 ID AAB82799 standard; protein; 232 AA.
 XX AAB82799;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
 XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Oryctolagus cuniculus.
 XX
 XX WO200164874-A2.
 PN
 XX 07-SEP-2001.
 PD

```

XX PF 28-FEB-2001; 2001WO-US006356.
XX PR 02-MAR-2000; 2000US-00517849.
XX PR 14-JUL-2000; 2000US-00616289.
XX PA (BOST-) BOSTON HEART FOUND INC.
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX DR WPI; 2001-565505/63.
XX DR N-PSDB; AAH26489.
XX PT New isolated low density lipoprotein binding polypeptide for treating,
XX PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX PS Claim 13(c); Fig 3; 143pp; English.
XX CC The present sequence is that of a partial sequence of novel rabbit low
XX CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
XX CC deduced from an isolated cDNA clone (see AAH26489). Full-length rabbit
XX CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed
XX CC polypeptides of the invention, termed LBPs, that are capable of binding
XX CC to native and methylated low density lipoproteins. Also claimed are
XX CC biologically active fragments and analogues of LBPs, polynucleotides
XX CC encoding LBPs, as well as expression vectors, cells and methods of
XX CC producing the LBPs. Methods of determining if an animal is at risk for
XX CC atherosclerosis, methods for evaluating an agent for use in treating
XX CC atherosclerosis, and methods for treating a cell having an abnormality in
XX CC structure or metabolism of LBP are also claimed, as are pharmaceutical
XX CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
XX CC compositions
XX SQ Sequence 232 AA;

Query Match 89.5%; Score 1047.5; DB 4; Length 232;
Best Local Similarity 88.6%; Pred. No. 2e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQQHQLN-GERGQSAKERVK 57
DB 13 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQQHQLN-GERGQSAKERVK 72
QY 58 EWTPCGPHQGGDEGRGAPAGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPKAL 117
DB 73 EWSLGGPHFGQEGRGAPAGSGTRQVFSMAALSKEGGSASTTGPDPSPSPVPLPGKPKAL 132
QY 118 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 177
DB 133 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 192
QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 193 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 232

RESULT 6
AAB82800
ID AAB82800 standard; protein; 252 AA.
XX AC AAB82800;
XX AC AAB82800;
XX DT 12-NOV-2001 (first entry)
XX DE Rabbit low density lipoprotein binding protein 3 (LBP-3).
XX DE Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
XX KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX OS Oryctolagus cuniculus.
XX OS WO200164874-A2.
XX PN
XX FH

Query Match 89.5%; Score 1047.5; DB 4; Length 252;
Best Local Similarity 88.6%; Pred. No. 2.2e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQQHQLN-GERGQSAKERVK 57
DB 33 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQQHQLN-GERGQSAKERVK 92
QY 58 EWTPCGPHQGGDEGRGAPAGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPKAL 117
DB 93 EWSLGGPHFGQEGRGAPAGSGTRQVFSMAALSKEGGSASTTGPDPSPSPVPLPGKPKAL 152
QY 118 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 177
DB 153 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 212
QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 213 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 252

RESULT 7
AAB82798
ID AAB82798 standard; protein; 317 AA.
XX AC AAB82798;
XX AC AAB82798;
XX DT 12-NOV-2001 (first entry)
XX DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
XX DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
XX KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX OS Oryctolagus cuniculus.
XX OS WO200164874-A2.
XX PN
XX FH

```

FT Misc-difference 10 /note= "encoded by TAG"

FT XX WO200164874-A2.
PN XX 07-SEP-2001.
PD XX
PF XX 28-FEB-2001; 2001WO-US006356.
PX XX 02-MAR-2000; 2000US-00517849.
PR XX 14-JUL-2000; 2000US-00616289.
PS XX (BOST-) BOSTON HEART FOUND INC.
PT XX
PX XX Lees AM, Lees RS, Law SW, Arjona AA;
PI XX WPI; 2001-565505/63.
DR XX N-PSDB; AAH26488.

XX New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX Claim 13(b); Fig 2b; 143pp; English.
PS
XX The present sequence is that of a partial sequence of novel rabbit low
CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
CC deduced from an isolated cDNA clone (see AAH26488). Full-length rabbit
CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed
CC polypeptides of the invention, termed LBPs, that are capable of binding
CC to native and methylated low density lipoproteins. Also claimed are
CC biologically active fragments and analogues of LBPs, polynucleotides
CC encoding LBPs, as well as expression vectors, cells and methods of
CC producing the LBPs. Methods of determining if an animal is at risk for
CC atherosclerosis, methods for evaluating an agent for use in treating
CC atherosclerosis, and methods for treating a cell having an abnormality in
CC structure or metabolism of LBP are also claimed, as are pharmaceutical
CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
CC compositions

XX SQ Sequence 317 AA;
Query Match 89.5%; Score 1047.5; DB 4; Length 317;
Best Local Similarity 88.6%; Pred. No. 2.9e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;
QY 1 EERVLEKEEEDDEDEDD--VSEGSVPESDRPAGAHQHLN-GERGPOSASERVK 57
DB 98 EERVLEKEEEDDEDEDD--VSEGSVPESDRPAGAHQHLN-GERGPOSASERVK 157
QY 58 EWTPCPHQDGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPAL 117
DB 158 ENSLCGPHQDGRGPAGSGTRQVFSMAALKEGTSASSTGPDSPSPVPLPPGKPAL 217
QY 118 PGADGTPFGCPGRKEKPSDPVEMTVMDVVEYFTEAGFPQATAFQEQEIDGKSLLMQR 177
DB 218 PGADGTPFGCPGRKEKPSDPVEMTVMDVVEYFTEAGFPQATAFQEQEIDGKSLLMQR 277
QY 178 TDVLTGLSIRLGPALKIYEHKIVLQOQHFFEDDDPDGFLG 217
DB 278 TDVLTGLSIRLGPALKIYEHKIVLQOQHFFEDDDPDGFLG 317

RESULT 8
AAB82807
ID AAB82807 standard; protein; 550 AA.
XX AC AAB82807;
XX AC AAB82807;
DT 12-NOV-2001 (first entry)
XX
DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
KW

KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX Oryctolagus cuniculus.
XX WO200164874-A2.
XX 07-SEP-2001.
XX 28-FEB-2001; 2001WO-US006356.
XX 02-MAR-2000; 2000US-00517849.
XX 14-JUL-2000; 2000US-00616289.
XX (BOST-) BOSTON HEART FOUND INC.
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX N-PSDB; AAH26500.
XX New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX Claim 13(k); Fig 2a; 143pp; English.
PS
XX The present sequence is that of a full-length sequence of novel rabbit
CC low density lipoprotein binding protein 2 (LBP-2). The amino acid
CC sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit
CC LBP-2 is an example of claimed polypeptides of the invention, termed
CC LBPs, that are capable of binding to native and methylated low density
CC lipoproteins. Also claimed are biologically active fragments and
CC analogues of LBPs, polynucleotides encoding LBPs, as well as expression
CC vectors, cells and methods of producing the LBPs. Polypeptides having
CC amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-
CC 18) of the present sequence are claimed. Methods of determining if an
CC animal is at risk for atherosclerosis, methods for evaluating an agent
CC for use in treating atherosclerosis, and methods for treating a cell
CC having an abnormality in structure or metabolism of LBP are also claimed,
CC as are pharmaceutical compositions comprising an LBP polypeptide or
CC nucleic acid, and vaccine compositions

XX SQ Sequence 550 AA;
Query Match 89.5%; Score 1047.5; DB 4; Length 550;
Best Local Similarity 88.6%; Pred. No. 5.7e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;
QY 1 EERVLEKEEEDDEDEDD--VSEGSVPESDRPAGAHQHLN-GERGPOSASERVK 57
DB 331 EERVLEKEEEDDEDEDD--VSEGSVPESDRPAGAHQHLN-GERGPOSASERVK 390
QY 58 EWTPCPHQDGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPAL 117
DB 391 ENSLCGPHQDGRGPAGSGTRQVFSMAALKEGTSASSTGPDSPSPVPLPPGKPAL 450
QY 118 PGADGTPFGCPGRKEKPSDPVEMTVMDVVEYFTEAGFPQATAFQEQEIDGKSLLMQR 177
DB 451 PGADGTPFGCPGRKEKPSDPVEMTVMDVVEYFTEAGFPQATAFQEQEIDGKSLLMQR 510
QY 178 TDVLTGLSIRLGPALKIYEHKIVLQOQHFFEDDDPDGFLG 217
DB 511 TDVLTGLSIRLGPALKIYEHKIVLQOQHFFEDDDPDGFLG 550

RESULT 9
AAW49038
ID AAW49038 standard; protein; 317 AA.
XX AC AAW49038;
XX AC AAW49038;
DT 09-NOV-1998 (first entry)
XX
DE Rabbit low density lipoprotein binding protein LBP-2.

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 132 AA;

Query Match 60.5%; Score 708; DB 7; Length 132;
 Best Local Similarity 100.0%; Pred. No. 2.5e-53;
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 MAAMNKEGGTASVATGSDSPVLPFGKALPGADGTPGCGPRKEKSPDVEVTVM 145
 DB 1 MAAMNKEGGTASVATGSDSPVLPFGKALPGADGTPGCGPRKEKSPDVEVTVM 60

QY 146 VVEYFTAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIVLQOG 205
 DB 61 VVEYFTAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIVLQOG 120

QY 206 HFEDDDPDGFLG 217
 DB 121 HFEDDDPDGFLG 132

RESULT 11

AAU31793
 ID AAU31793 standard; protein; 316 AA.

XX AC AAU31793;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #2284.

XX KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN W0200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic

XX PS vaccination, testing and therapy.

XX PS Claim 20; Page 509; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy. The proteins
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or

CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention

XX SQ Sequence 316 AA;

Query Match 29.9%; Score 350; DB 4; Length 316;
 Best Local Similarity 38.7%; Pred. No. 7e-22;
 Matches 103; Conservative 7; Mismatches 68; Indels 98; Gaps 12;

QY 1 EERVLEKEEEDD-DEDEDEDDVSEGSEV-PESDRPAGAHQHLNGERCPOQSAKERVKE 58
 DB 53 EERVLEKEEEDD-DEDEDEDDVSEGSEV-PESDRPAGAHQHLNGERCPOQSAKERVKE 98

QY 59 WTPCG----PHQQDEG-RGPAPGSGTRQVFSMAAMNKEGGT-----A 96
 DB 99 -TASGDLRVPRRSRSGPPTAFGPGXKAGASPGQRHPPGCSQWQXTRKGEQVFFSPPA 157

QY 97 SVATGSDSPSPVLPFGKALPGADGTPGCGPRKEKSPD-VEVTVM DV-VYFTFTEAG 154
 DB 158 SVATGSDSPSPVLPFGKALPGADGTPGCGPRKEKSPD-VEVTVM DV-VYFTFTEAG 217

QY 155 PPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHK----- 198
 DB 218 PPEQATAF-----RAGNKKWQIFAFAAAHRCASPACPSPRAS 254

QY 199 ---IKVLQGHFEDD-----DPDGFL 216
 DB 255 PENLRAPHQAGASARPLXGMMDPDGFL 280

RESULT 12

AAAB34365

ID AAB34365 standard; protein; 61 AA.

XX AC AAB34365;

XX DT 26-JAN-2001 (first entry)

XX DE Human secreted protein sequence encoded by gene 9 SEQ ID NO:126.

XX KW Human; secreted protein; diagnosis; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active general; vulnary;
 KW gastrointestinal; nephrotropic; antinfecive; gynaecological;
 KW and antibacterial; gene therapy; detection; cancer; chromosome marker;
 KW chromosome identification; neural disorder; immune disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW proliferative disorder; wound healing; infectious disease; preservative;
 KW food additive.

XX OS Homo sapiens.

XX PN W0200056883-A1.

XX PD 28-SEP-2000.

XX PF 16-MAR-2000; 2000WO-US006822.

XX PR 23-MAR-1999; 99US-0126054P.

XX PR 10-DEC-1999; 99US-0169916P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-587666/55.

XX CC Human secreted proteins and gene sequences encoding them, useful for
 CC detecting, preventing, and treating disorders such as cancer,
 CC neurological disorders and immune system disorders.
 XX Disclosure; Page 396-397; 429pp; English.

XX The polynucleotide sequences given in AAC59566 to AAC59614 encode the
 CC human secreted proteins given in AAB34299 to AAB34347. AAB34348 to
 CC AAB34437 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular
 CC active general; vulnary; gastrointestinal; nephrotropic; antiinfective;
 CC gynaecological; and antibacterial. The polynucleotides can be used for
 CC the detection of various disorders such as cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The secreted proteins can be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wound healing, and infectious diseases. The proteins can also
 CC be used as a food additive or preservative to increase or decrease
 CC storage capabilities. AAC59557 to AAC59565 and AAB34298 represent
 CC sequences used in the exemplification of the present invention

XX Sequence 61 AA;
 SQ

Query Match 26.7%; Score 312; DB 3; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.9e-19;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 PSDPVEVTVMVVVEYFTEAGFPQATAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKI 194
 DB 1 PSDPVEVTVMVVVEYFTEAGFPQATAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKI 60
 QY 195 Y 195
 DB 61 Y 61

RESULT 13
 ADA55289
 ID ADA55289 standard; protein; 116 AA.
 AC ADA55289;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 DE Human protein, SEQ ID 2857.
 XX
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.
 XX
 OS Homo sapiens.
 XX
 XX EP1293569-A2.
 PN
 XX 19-MAR-2003.
 PD
 XX 21-MAR-2002; 2002EP-00006586.
 PF
 XX 14-SEP-2001; 2001JP-00328381.
 PR
 XX 24-JAN-2002; 2002US-0350435P.
 PR
 XX (HELI-) HELIX RES INST.
 PA
 PA (HEAS-) RES ASSOC BIO TECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 XX WPI; 2003-395539/38.
 DR
 DR N-PSDB; ADA53650.
 XX
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 14; SEQ ID NO 2857; 205pp; English.
 PS
 XX The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 116 AA;
 SQ

Query Match 24.4%; Score 285.5; DB 6; Length 116;
 Best Local Similarity 46.9%; Pred. No. 8e-17;
 Matches 60; Conservative 14; Mismatches 26; Indels 28; Gaps 2;

QY 83 VFSMAAMNKEGTASVATGDSFSPVPLPPGKALPGADGTGFGCPGKRPSPDVEWT 142
 DB 15 MLSDVMENKENGSGVGKNSMEN-----GR---PPDPADWA 46
 QY 143 VMDVVVEYFTEAGFPQATAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKIYEHKVL 202
 DB 47 VMDVVVYFRTVGFEQASAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKIYEHKPL 106
 QY 203 QQGHFEDD 210
 DB 107 QYKHLKNV 114

RESULT 14
 AAM39592
 ID AAM39592 standard; protein; 102 AA.
 XX
 AC AAM39592;
 XX
 XX 22-OCT-2001 (first entry)
 DT
 DE Human polypeptide SEQ ID NO 2737.
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200153312-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX 26-DEC-2000; 2000WO-US034263.
 PF
 XX 23-DEC-1999; 99US-00471275.
 PR
 XX 21-JAN-2000; 2000US-00488725.
 PR
 XX 25-APR-2000; 2000US-00552317.
 PR
 XX 20-JUN-2000; 2000US-00598042.
 PR
 XX 19-JUL-2000; 2000US-00620312.
 PR
 XX 03-AUG-2000; 2000US-00653450.
 PR
 XX 14-SEP-2000; 2000US-00662191.
 PR
 XX 19-OCT-2000; 2000US-00691036.
 PR
 XX 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI; 2001-442253/47.
 DR
 DR N-PSDB; AAI58748.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX Example 4; SEQ ID NO 2737; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA138642-AA142213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 XX part of the printed specification

SQ Sequence 102 AA;

Query Match 24.4%; Score 285; DB 4; Length 102;

Best Local Similarity 46.9%; Pred. No. 7.5e-17;
 Matches 60; Conservative 14; Mismatches 26; Indels 28; Gaps 2;

QY 83 VFSMAAMNKEGGTASVATGPDSPVLPKPKALPGADGTPFCGPKRKEKPSDPVEWT 142

DB 1 MLSVDMENKENGSGVGVNSMEN-----GR---PPDPADWA 32

QY 143 VMDVVYFTAGPPEQATAFQEQEIDGKSLLMORTDVLTLGLSIRLGPALKIYEHKVL 202

DB 33 VMDVVNYFRTVGPEEQASAFQEQEIDGKSLLMTRNDVLTGLQLKLGPAKLIYEHVKPL 92

QY 203 QQGHFEDD 210

DB 93 QTKHLKNN 100

RESULT 15

ABB50215
 ID ABB50215 standard; protein; 102 AA.

AC ABB50215;

DT 05-FEB-2002 (first entry)

DE Human transcription factor TRFX-66.

XX Human; transcription factor; TRFX; cell proliferative disease;

KW autoimmune disease; inflammation; neurological disease;

KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;

XX neuroprotective; antiinflammatory; gene therapy.

OS Homo sapiens.

XX WO200172777-A2.

XX 04-OCT-2001.

XX 13-MAR-2001; 2001WO-US008117.

XX 13-MAR-2000; 2000US-0188986P.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;

PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;

PI Reddy R;

XX WPI; 2001-570896/64.

DR N-PSDB; ABA83039.

XX

PT Novel transcription factor polypeptides, used to treat diseases
 PT associated with altered activity and expression of TRFX, and to screen
 PT for agents capable of modulating its activity.

XX Claim 1; Page 210; 327pp; English.

XX The present sequence is the protein sequence for a human transcription
 CC factor. The transcription factor and its coding sequence are useful in
 CC the diagnosis, treatment and prevention of diseases associated with
 CC altered expression of the transcription factor e.g. cell proliferative,
 CC autoimmune/inflammatory, neurological and developmental disorders. A
 CC number of specific disorders/diseases are given in the specification,
 CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
 CC allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic
 CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
 CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
 CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
 CC colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
 CC disease, stroke, and viral, bacterial, fungal and protozoal infections

XX SQ Sequence 102 AA;

Query Match 24.4%; Score 285; DB 4; Length 102;

Best Local Similarity 46.9%; Pred. No. 7.5e-17;
 Matches 60; Conservative 14; Mismatches 26; Indels 28; Gaps 2;

QY 83 VFSMAAMNKEGGTASVATGPDSPVLPKPKALPGADGTPFCGPKRKEKPSDPVEWT 142

DB 1 MLSVDMENKENGSGVGVNSMEN-----GR---PPDPADWA 32

QY 143 VMDVVYFTAGPPEQATAFQEQEIDGKSLLMORTDVLTLGLSIRLGPALKIYEHKVL 202

DB 33 VMDVVNYFRTVGPEEQASAFQEQEIDGKSLLMTRNDVLTGLQLKLGPAKLIYEHVKPL 92

QY 203 QQGHFEDD 210

DB 93 QTKHLKNN 100

Search completed: September 20, 2005, 12:34:40

Job time : 131.084 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 32.9351 Seconds
(without alignments)
491.841 Million cell updates/sec

Title: US-10-671-242-7

Perfect score: 1170

Sequence: 1 BERVLSKBEEDDEDEDE.....HIKVLQGHFDDDPDGLG 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1170	100.0	217	3	US-08-979-608A-7
2	1170	100.0	217	4	US-09-517-849-7
3	1170	100.0	217	4	US-09-616-289-7
4	1170	100.0	538	4	US-09-616-289-7
5	1047.5	89.5	232	3	US-08-979-608A-3
6	1047.5	89.5	232	4	US-09-517-849-3
7	1047.5	89.5	232	4	US-09-616-289-3
8	1047.5	89.5	252	3	US-08-979-608A-4
9	1047.5	89.5	252	4	US-09-517-849-4
10	1047.5	89.5	252	4	US-09-616-289-4
11	1047.5	89.5	317	3	US-08-979-608A-2
12	1047.5	89.5	317	4	US-09-517-849-2
13	1047.5	89.5	317	4	US-09-616-289-2
14	1047.5	89.5	550	4	US-09-616-289-47
15	153	13.1	1004	3	US-08-916-352-2
16	153	13.1	1072	4	US-09-949-016-8072
17	149	12.7	1004	4	US-09-949-016-6496
18	146.5	12.5	231	3	US-08-974-380-2
19	146.5	12.5	231	4	US-09-546-977A-2
20	146.5	12.5	231	4	US-09-654-466-2
21	146.5	12.5	340	4	US-09-949-016-10568
22	142.5	12.2	577	2	US-08-852-153-4
23	142.5	12.2	591	2	US-08-852-153-6
24	142.5	12.2	620	2	US-08-852-153-2
25	141.5	12.1	654	2	US-08-852-153-8
26	136	11.6	26	3	US-08-979-608A-20
27	136	11.6	26	4	US-09-517-849-20

28	136	11.6	26	4	US-09-616-289-20	Sequence 20, Appl
29	132	11.3	797	4	US-09-949-016-9676	Sequence 9676, Ap
30	133	10.5	754	4	US-09-392-714-20	Sequence 20, Appl
31	123	10.5	801	4	US-09-949-016-6588	Sequence 6588, Ap
32	123	10.5	810	4	US-09-949-016-11152	Sequence 11152, A
33	123	10.5	905	2	US-08-574-959A-9	Sequence 9, Appl1
34	123	10.5	905	3	US-09-357-014-9	Sequence 9, Appl1
35	123	10.5	1135	2	US-08-574-959A-7	Sequence 7, Appl1
36	123	10.5	1135	3	US-09-357-014-7	Sequence 7, Appl1
37	119.5	10.2	478	3	US-08-155-888-2	Sequence 2, Appl1
38	118.5	10.1	247	4	US-09-538-092-890	Sequence 890, App
39	118.5	10.1	258	4	US-09-949-016-9409	Sequence 9409, Ap
40	118	10.1	1341	4	US-09-949-016-6890	Sequence 6890, Ap
41	118	10.1	1344	4	US-09-949-016-10925	Sequence 10925, A
42	115.5	9.9	433	4	US-09-949-016-6497	Sequence 6497, Ap
43	115	9.8	2451	4	US-09-949-016-9675	Sequence 9675, Ap
44	114.5	9.8	1505	4	US-09-538-092-1102	Sequence 1102, Ap
45	114	9.7	411	4	US-09-252-991A-24632	Sequence 24632, A

ALIGNMENTS

RESULT 1

US-08-979-608A-7
; Sequence 7, Application US/08979608A
; Patent No. 6355451

GENERAL INFORMATION:

APPLICANT: Lees, Ann M. S.

Law, Robert S.

Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,608A

FILING DATE: 26-Nov. 6355451-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,547

FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930

FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 217 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-08-979-608A-7

Query Match 100.0%; Score 1170; DB 3; Length 217;

APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIOR FILING DATE: 2000-07-14
PRIOR FILING DATE: 2000-03-02
PRIOR FILING DATE: 1997-11-26
PRIOR FILING DATE: 1996-11-27
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 538
TYPE: PRT
ORGANISM: Homo sapiens
US-09-616-289-43

Query Match 100.0%; Score 1170; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 9.1e-97;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPSQAKERVKEWT 60
DB 322 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPSQAKERVKEWT 381
QY 61 PCGPHQODRGCPAGCGTRQVFSMAAMNKEGTASVATGPDSPVPLPGKPALPGA 120
DB 382 PCGPHQODRGCPAGCGTRQVFSMAAMNKEGTASVATGPDSPVPLPGKPALPGA 441
QY 121 DGTFFGCPGCKPKSDPPVETVMDVVEYFTEAGFPEQATAFOEQEIDGKSLLMQRTDV 180
DB 442 DGTFFGCPGCKPKSDPPVETVMDVVEYFTEAGFPEQATAFOEQEIDGKSLLMQRTDV 501
QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 5
US-08-979-608A-3
Sequence 3, Application US/08979608A
Patent No. 6355451
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,608A

FILING DATE: 26-No. 6355451-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,547
FILING DATE: 03-JUN-1997
APPLICATION NUMBER: US 60/031,930
FILING DATE: 27-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match 89.5%; Score 1047.5; DB 3; Length 232;
Best Local Similarity 88.6%; Pred. No. 3.1e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLN-GERGPOSQAKERVK 57
DB 13 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPGQTAKERAK 72
QY 58 EWTCPGPHQODRGCPAGCGTRQVFSMAAMNKEGTASVATGPDSPVPLPGKPAL 117
DB 73 ENSLGCSPGPHQODRGCPAGCGTRQVFSMAALKEGSSASTTGPDSPPVPLPGKPAL 132
QY 118 PGADGTPGCPGCKPKSDPPVETVMDVVEYFTEAGFPEQATAFOEQEIDGKSLLMQ 177
DB 133 PGADGTPGCPGCKPKSDPPVETVMDVVEYFTEAGFPEQATAFOEQEIDGKSLLMQ 192
QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 193 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 232

RESULT 6
US-09-517-849-3
Sequence 3, Application US/09517849
Patent No. 6605588
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608

```

; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-517-849-3

Query Match      89.5%; Score 1047.5; DB 4; Length 232;
Best Local Similarity 88.6%; Pred. No. 3.1e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEEEDEDEDD--VSEGEVPESDRPAGAQHHQLN-GERGPOS AKERVK 57
Db 13 EERVLEKEEEEEDEDDDDVVSEGEVPESDRPAGAQHHQLNCGERGPGTAKERAK 72

QY 58 EWTGCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPAL 117
Db 73 EWSLGGPHGQEGRGPAAGSGTQVFSMAALSKEGGSASTTGPDSPSPVPLPGKPAL 132

QY 118 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGFPPEQATAFQEQEIDGKSLMLMOR 177
Db 133 PGADGTPFGCPGRKEKPADPVETWMDVVEYFTAGFPPEQATAFQEQEIDGKSLMLMOR 192

QY 178 TDVLTGLSIRLGPALKIYEHHIKVLQOQHGFEDDDPDGFLG 217
Db 193 TDVLTGLSIRLGPALKIYEHHIKVLQOQHGFEDDDPDGFLG 232

RESULT 8
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match      89.5%; Score 1047.5; DB 3; Length 252;
Best Local Similarity 88.6%; Pred. No. 3.4e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEEEDEDEDD--VSEGEVPESDRPAGAQHHQLN-GERGPOS AKERVK 57
Db 33 EERVLEKEEEEEDEDDDDVVSEGEVPESDRPAGAQHHQLNCGERGPGTAKERAK 92

```

Qy	58	EWTPCGPHQGDEGRGPAGSGSTRQVFSMAANNKEGGTASVARGDPSDPSPVPLPPGKPAL	117
Db	93	EWSLCGPHPGQEGRGPAAGSCTRQVFSMAALSKEGGSASTTGDPDSPSPVPLPPGKPAL	152
Qy	118	PGADGTTPGCPPGRKEKESDPVEWTVMDVVVFTEAGFPPEQTATAPOEQBIDGKSLLLMQR	177
Db	153	PGADGTTPGCAPGRKEKPDPVEWTVMDVVVFTEAGFPPEQTATAPOEQIDGKSLLLMQR	212
Qy	178	TDVLTLGLSIRLGPALKIYEHHIKVLOQHFFEDDDPDGFLG	217
Db	213	TDVLTLGLSIRLGPALKIYEHHIKVLOQHFFEDDDPEGFLG	252

```

RESULT 9
US-09-517-849-4
: Sequence 4, Application US/09517849
: Patent No. 6605588
: GENERAL INFORMATION:
: APPLICANT: Lees, Ann M.
: Lees, Robert S.
: Law, Simon W.
: Arjona, Anibal A.
: TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
: TREATING ATHEROSCLEROSIS
: NUMBER OF SEQUENCES: 42

```

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fast-SEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-517-849-4

Query Match	89.5%;	Score	1047.5;	DB	4;	Length	252;
Best Local Similarity	88.6%;	Pred. No.	3.4e-86;				
Matches	195;	Conservative	15;	Mismatches	7;	Indels	3;
Gaps	2;						

Qy	1	B	E	R	V	L	K	E	E	E	D	E	D	E	D	--	V	S	G	S	E	V	P	S	D	R	P	A	G	A	H	O	L	N	-	G	E	R	G	P	S	A	K	R	V	K	57					
Db	33	E	E	R	V	L	K	E	E	E	E	D	E	D	D	D	D	V	V	S	G	S	E	V	P	S	D	R	P	A	G	A	H	O	L	N	G	E	R	G	P	T	A	K	E	R	A	K	92			
Qy	58	E	W	T	P	C	G	H	O	C	O	D	E	R	G	P	A	G	S	G	T	Q	V	F	S	M	A	M	N	K	E	G	T	A	S	V	A	T	C	P	D	S	P	P	L	P	P	G	K	A	L	117
Db	93	E	W	S	I	C	G	P	H	P	G	O	E	R	G	P	A	G	S	G	T	Q	V	F	S	M	A	L	S	K	E	G	S	A	S	T	T	C	P	D	S	P	P	L	P	P	G	K	A	L	152	

```

Qy 118 PGADGTPGCGPRKEKSDPVWMTVDVVEYFTTAGEPEQATAQEQEIDGKSULLMQR 177
Db 153 PGADGTPGCGPAGRKEKADPVWMTVDVVEYFTTAGEPEQATAQEQEIDGKSULLMQR 212
Qy 178 TDVLTGLSIRLGPALKIYEHHIKVLQOQHFEDDDDPDGFLG 217
Db 213 TDVLTGLSIRLGPALKIYEHHIKVLQOQHFEDDDDPEGFLG 252

RESULT 10
US-09-616-289-4
; Sequence 4, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-4

```

Query Match	89.5%; Score 1047.5; DB 4; Length 252;
Best Local Similarity	88.6%; Pred. No. 3.4e-86;
Matches 195; Conservative	15; Mismatches 7; Indels 3; Gaps 2;
Qy	1 EERVLEKXEEDDDDEDED--VSGSEVPESDRPAGAQHQHQLN--GERGPQSAAKERVK 57
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	33 EERVLEKXEEDDEDDDDVVSGSEVPESDRPAGAQHQHQLNGGRCGPQTAKERAK 92
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	58 EWTFCGHQGDGGRGAPAGSTGTQVFSSMAANKEGGTASVATGPDSPSPVPLPPGPKPAL 117
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	93 EWSLCGPHPGQEGRGPAAGSTGTQVFSSMAALKEGGSASTTTPDPSPVPLPPGPKPAL 152
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	118 PGADGTFPGCPPGRKRKSPDPVEVTWMDVVEYFTPEAGFPQCATAFQGEIDGKSLLLMOR 177
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	153 PGADGTFPGCPAGRKEPADPVEVTWMDVVEYFTPEAGFPQCATAFQGEIDGKSLLLMOR 212
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	178 TDVLVTGLSIRLGPALKIYEHHIKVLOQGHFEDDDPDGFLG 217
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Qy	213 TDVLVTGLSIRLGPALKIYEHHIKVLOQGHFEDDDPDGFLG 252
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::

RESULT 11
US-08-979-608A-2
; Sequence 2, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Iaw, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS

QY 1 EERVLEKEEEEEDEDEDEDD--VSEGEVPESEDRPAGAHQHQLN-GERGPQSAKERVK 57

Search completed: September 20, 2005, 12:45:28
Job time : 33.9351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 119.788 Seconds
(without alignments)
733.538 Million cell updates/sec

Title: US-10-671-242-7
Perfect score: 1170
Sequence: 1 BERVLEKKEEEDDEDEEE.....HIKVLQQGHFEDDDPGFLG 217

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues
Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.PEP.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.PEP.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.PEP.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.PEP.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.PEP.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.PEP.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.PEP.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.PEP.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.PEP.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.PEP.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.PEP.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.PEP.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.PEP.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.PEP.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.PEP.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.PEP.*
17:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.PEP.*
18:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.PEP.*
19:	/cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.PEP.*
20:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.PEP.*
21:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.PEP.*
22:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1170	100.0	217	9	US-09-962-055-7 Sequence 7, Appli
2	1170	100.0	217	9	US-09-976-740-7 Sequence 7, Appli
3	1170	100.0	217	13	US-10-023-529-7 Sequence 7, Appli
4	1170	100.0	217	13	US-10-023-523-7 Sequence 7, Appli
5	1170	100.0	217	15	US-10-616-187-7 Sequence 7, Appli
6	1170	100.0	217	15	US-10-671-242-7 Sequence 7, Appli
7	1170	100.0	538	9	US-09-976-740-43 Sequence 43, Appl
8	1170	100.0	538	13	US-10-023-529-43 Sequence 43, Appl
9	1170	100.0	538	13	US-10-023-523-43 Sequence 43, Appl
10	1170	100.0	538	15	US-10-616-187-43 Sequence 43, Appl
11	1170	100.0	538	15	US-10-671-242-43 Sequence 43, Appl

12	1161	99.2	241	9	US-09-925-298-665 Sequence 665, App
13	1161	99.2	241	14	US-10-102-806-665 Sequence 665, App
14	1047.5	89.5	232	9	US-09-962-055-3 Sequence 3, Appli
15	1047.5	89.5	232	9	US-09-976-740-3 Sequence 3, Appli
16	1047.5	89.5	232	13	US-10-023-529-3 Sequence 3, Appli
17	1047.5	89.5	232	13	US-10-023-523-3 Sequence 3, Appli
18	1047.5	89.5	232	15	US-10-616-187-3 Sequence 3, Appli
19	1047.5	89.5	232	15	US-10-671-242-3 Sequence 3, Appli
20	1047.5	89.5	252	9	US-09-962-055-4 Sequence 4, Appli
21	1047.5	89.5	252	9	US-09-976-740-4 Sequence 4, Appli
22	1047.5	89.5	252	13	US-10-023-529-4 Sequence 4, Appli
23	1047.5	89.5	252	13	US-10-023-523-4 Sequence 4, Appli
24	1047.5	89.5	252	15	US-10-616-187-4 Sequence 4, Appli
25	1047.5	89.5	252	15	US-10-671-242-4 Sequence 4, Appli
26	1047.5	89.5	317	9	US-09-962-055-2 Sequence 2, Appli
27	1047.5	89.5	317	9	US-09-976-740-2 Sequence 2, Appli
28	1047.5	89.5	317	13	US-10-023-529-2 Sequence 2, Appli
29	1047.5	89.5	317	13	US-10-023-523-2 Sequence 2, Appli
30	1047.5	89.5	317	15	US-10-616-187-2 Sequence 2, Appli
31	1047.5	89.5	317	15	US-10-671-242-2 Sequence 2, Appli
32	1047.5	89.5	550	9	US-09-976-740-47 Sequence 47, Appl
33	1047.5	89.5	550	13	US-10-023-529-47 Sequence 47, Appl
34	1047.5	89.5	550	13	US-10-023-523-47 Sequence 47, Appl
35	1047.5	89.5	550	15	US-10-616-187-47 Sequence 47, Appl
36	1047.5	89.5	550	15	US-10-671-242-47 Sequence 47, Appl
37	285.5	24.4	116	15	US-10-094-749-2857 Sequence 2857, Ap
38	285	24.4	102	15	US-10-221-625-66 Sequence 66, Appl
39	169	14.4	319	14	US-10-106-698-4861 Sequence 4861, Ap
40	166.5	14.2	894	16	US-10-713-993-647 Sequence 647, App
41	166.5	14.2	894	17	US-10-485-555-16 Sequence 16, Appl
42	153.5	13.1	356	15	US-10-104-047-3703 Sequence 3703, Ap
43	151.5	12.9	193	18	US-10-450-763-51392 Sequence 51392, A
44	151	12.9	1012	18	US-10-953-460-12 Sequence 12, Appl
45	149	12.7	1004	16	US-10-723-860-772 Sequence 772, App

ALIGNMENTS

RESULT 1
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; City: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/09/962,055
; APPLICATION DATA:
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match      100.0%; Score 1170; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
DB 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
QY 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
QY 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
DB 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 2
US-09-976-740-7
; Sequence 7, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-7

Query Match      100.0%; Score 1170; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
DB 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
QY 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
QY 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
DB 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 2
US-09-976-740-7
; Sequence 7, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-7
```

```
Db 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
QY 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
QY 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
DB 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 3
US-10-023-529-7
; Sequence 7, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 1170; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
DB 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAQHHQHLNGERGQPSAKERVKEWT 60
QY 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
QY 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
DB 121 DGTFFGCPGGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMORTDV 180
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 4
US-10-023-529-7
; Sequence 7, Application US/10023523
; Publication No. US20020152485A1
```

```

; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-7

Query Match      100.0%; Score 1170; DB 15; Length 217;
Best Local Similarity .100.0%; Pred. No. 1.8e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 EERVLEKEEEDDEDEDDYSEGSEVPESDRPAGAQHHQLNGERGPPQSAKERVKEWT 60
Db 1 EERVLEKEEEDDEDEDDYSEGSEVPESDRPAGAQHHQLNGERGPPQSAKERVKEWT 60

Qy 61 PCGPHQODGRGPAGSGTRQVFSMAANNKEGGTASVATGPDSPVPPLPPGKPALPGA 120
Db 61 PCGPHQODGRGPAGSGTRQVFSMAANNKEGGTASVATGPDSPVPPLPPGKPALPGA 120

Qy 121 DGPFPFGCPGRKCKPSPDVSWTVMVYVYFTTEAGFPQQAATAFQOEIDGKSLLLMORTDV 180
Db 121 DGPFPFGCPGRKCKPSPDVSWTVMVYVYFTTEAGFPQQAATAFQOEIDGKSLLLMORTDV 180

Qy 181 LTGLSIRLGPAALKIYEHHIKVLQOGHFEDDDPDGFLG 217
Db 181 LTGLSIRLGPAALKIYEHHIKVLQOGHFEDDDPDGFLG 217

RESULT 6
US-10-671-242-7
; Sequence 7, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 1170; DB 15; Length 217;
Best Local Similarity .100.0%; Pred. No. 1.8e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 EERVLEKEEEDDEDEDDYSEGSEVPESDRPAGAQHHQLNGERGPPQSAKERVKEWT 60
Db 1 EERVLEKEEEDDEDEDDYSEGSEVPESDRPAGAQHHQLNGERGPPQSAKERVKEWT 60

Qy 61 PCGPHQODGRGPAGSGTRQVFSMAANNKEGGTASVATGPDSPVPPLPPGKPALPGA 120

```

Db 61 PCGPHQDQEGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPPGKPALPGA 120
QY 121 DGTFFGCGPKRKEKPSDPEWTVMDVVEYFTEAGFPEQATAFQOEIDGKSLMLMORTDV 180
Db 121 DGTFFGCGPKRKEKPSDPEWTVMDVVEYFTEAGFPEQATAFQOEIDGKSLMLMORTDV 180
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217
Db 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217

RESULT 7

US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-43

Query Match 100.0%; Score 1170; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EERVLEKEEEDDEDEDEDDVSEGSVEPSDRPAGAHQHLNGERGPQSAKERVKEWT 60
Db 322 EERVLEKEEEDDEDEDEDDVSEGSVEPSDRPAGAHQHLNGERGPQSAKERVKEWT 381
QY 61 PCGPHQDQEGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPPGKPALPGA 120
Db 382 PCGPHQDQEGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPPGKPALPGA 441
QY 121 DGTFFGCGPKRKEKPSDPEWTVMDVVEYFTEAGFPEQATAFQOEIDGKSLMLMORTDV 180
Db 442 DGTFFGCGPKRKEKPSDPEWTVMDVVEYFTEAGFPEQATAFQOEIDGKSLMLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217
Db 502 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 538

RESULT 8

US-10-023-529-43
; Sequence 43, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-03-02
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-43
Query Match 100.0%; Score 1170; DB 13; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EERVLEKEEEDDEDEDEDDVSEGSVEPSDRPAGAHQHLNGERGPQSAKERVKEWT 60
Db 322 EERVLEKEEEDDEDEDEDDVSEGSVEPSDRPAGAHQHLNGERGPQSAKERVKEWT 381
QY 61 PCGPHQDQEGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPPGKPALPGA 120
Db 382 PCGPHQDQEGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPPGKPALPGA 441
QY 121 DGTFFGCGPKRKEKPSDPEWTVMDVVEYFTEAGFPEQATAFQOEIDGKSLMLMORTDV 180
Db 442 DGTFFGCGPKRKEKPSDPEWTVMDVVEYFTEAGFPEQATAFQOEIDGKSLMLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217
Db 502 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 538

RESULT 9
US-10-023-523-43
; Sequence 43, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-03-02
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43

```
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-43

Query Match      100.0%; Score 1170; DB 13; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 60
DB 322 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 381
QY 61 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 382 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 441
QY 121 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 180
DB 442 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538

RESULT 10
US-10-616-187-43
; Sequence 43, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-43

Query Match      100.0%; Score 1170; DB 15; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 60
DB 322 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 381
QY 61 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 382 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 441
QY 121 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 180
DB 442 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538

RESULT 11
US-10-671-242-43
; Sequence 43, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-43

Query Match      100.0%; Score 1170; DB 15; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 60
DB 322 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 381
QY 61 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 382 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 441
QY 121 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 180
DB 442 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538

RESULT 12
US-09-925-298-665
; Sequence 665, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
```

```
DB 442 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538

RESULT 11
US-10-671-242-43
; Sequence 43, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US/08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US/60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US/60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-43

Query Match      100.0%; Score 1170; DB 15; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 60
DB 322 EERVLEKEEEDDEDEDDVSEGEVPESDRPAGAHQHLNGERGPOSASAKERVKEWT 381
QY 61 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
DB 382 PCGPHQGDGRGPAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 441
QY 121 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 180
DB 442 DGTFFGCPGRKEKPSDPVETVMDVVVEYFTEAGPPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538

RESULT 12
US-09-925-298-665
; Sequence 665, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
```

/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 846
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 665
/ LENGTH: 241
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (9)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (122)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-665

Query Match 99.2%; Score 1161; DB 9; Length 241;
Best Local Similarity 99.1%; Pred. No. 9.7e-80;
Matches 215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPQSAKERVKWT 60
DB 25 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPQSAKERVKWT 84
QY 61 PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120
DB 85 PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 144
QY 121 DGTFFGCPGRKEKPSDPVEMTMDVVEYFTAGFPEQATAFQEQIDGKSLLMQRTDV 180
DB 145 DGTFFGCPGRKEKPSDPVEMTMDVVEYFTAGFPEQATAFQEQIDGKSLLMQRTDV 204
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217
DB 205 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 241

RESULT 13
US-10-102-806-665
/ Sequence 665, Application US/10102806
/ Publication No. US20030054421A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA103PIC1
/ CURRENT APPLICATION NUMBER: US/10/102,806
/ PRIOR FILING DATE: 2002-03-22
/ PRIOR APPLICATION NUMBER: 09/925,298
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05881
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 846
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 665
/ LENGTH: 241
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (9)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (122)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-665

Query Match 99.2%; Score 1161; DB 14; Length 241;
Best Local Similarity 99.1%; Pred. No. 9.7e-80;
Matches 215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPQSAKERVKWT 60
DB 25 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPQSAKERVKWT 84
QY 61 PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120
DB 85 PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 144
QY 121 DGTFFGCPGRKEKPSDPVEMTMDVVEYFTAGFPEQATAFQEQIDGKSLLMQRTDV 180
DB 145 DGTFFGCPGRKEKPSDPVEMTMDVVEYFTAGFPEQATAFQEQIDGKSLLMQRTDV 204
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217
DB 205 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 241

RESULT 14
US-09-962-055-3
/ Sequence 3, Application US/09962055
/ Patent No. US20020052033A1
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ Lees, Robert S.
/ Law, Simon W.
/ Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
/ TREATING ATHEROSCLEROSIS
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/962,055
/ FILING DATE: 24-Sep-2001
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/979,608
/ FILING DATE: 26-NOV-1997
/ APPLICATION NUMBER: US 60/031,930
/ FILING DATE: 27-NOV-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 232 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-962-055-3

Query Match 89.5%; Score 1047.5; DB 9; Length 232;
Best Local Similarity 88.6%; Pred. No. 3.5e-71;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;
QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPQSAKERVKWT 57
DB 13 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPQSAKERVKWT 72

```
QY 58 EWTCPGPHQDGRGPAGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPAL 117
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 EWSLCGPHPGQEGRGPAAGSGTRQVFSMAALSKEGGSASSTTGGPDSPSPVPLPPGKPAL 132
QY 118 PGADGTFPGCPGRKEKPSDPVETVMDVVVEYTEAGFPQATAFQEQEIDGKSLLIMOR 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 PGADGTFPGCPGRKEKPADPVETVMDVVVEYTEAGFPQATAFQEQEIDGKSLLIMOR 192
QY 178 TDVLTGLSIRLGPALKIYEHKIVLQOGHPFEDDDPDGFLG 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 193 TDVLTGLSIRLGPALKIYEHKIVLQOGHPFEDDDPDGFLG 232
```

```
RESULT 15
US-09-976-740-3
; Sequence 3, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-976-740-3
```

```
Query Match 89.5%; Score 1047.5; DB 9; Length 232;
Best Local Similarity 88.6%; Pred. No. 3.5e-71;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEDDEDEDEDD--VSEGSVPESDRPAGAHQHLN-GERGPOS AKERVK 57
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 13 EERVLEKEEEDDEDEDEDDDDVSEGSVPESDRPAGAHQHLNNGERGPTAKERAK 72
QY 58 EWTCPGPHQDGRGPAGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPAL 117
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 EWSLCGPHPGQEGRGPAAGSGTRQVFSMAALSKEGGSASSTTGGPDSPSPVPLPPGKPAL 132
QY 118 PGADGTFPGCPGRKEKPSDPVETVMDVVVEYTEAGFPQATAFQEQEIDGKSLLIMOR 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 PGADGTFPGCPGRKEKPADPVETVMDVVVEYTEAGFPQATAFQEQEIDGKSLLIMOR 192
QY 178 TDVLTGLSIRLGPALKIYEHKIVLQOGHPFEDDDPDGFLG 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 193 TDVLTGLSIRLGPALKIYEHKIVLQOGHPFEDDDPDGFLG 232
```

Search completed: September 20, 2005, 12:53:05
Job time : 120.788 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 110.226 Seconds
(without alignments)
1008.117 Million cell updates/sec

Title: US-10-671-242-7

Perfect score: 1170

Sequence: 1 BERVLSKBEEDDEDEDE.....HIKVLQGHFEDDDPDLG 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1170	100.0	236	2 Q96IM4	Q96IM4 homo sapien
2	1170	100.0	285	2 Q6P0R3	Q6P0R3 homo sapien
3	1170	100.0	295	2 Q6PT57	Q6PT57 homo sapien
4	1170	100.0	538	2 Q6SPF0	Q6SPF0 homo sapien
5	1047.5	89.5	550	2 Q6SPF9	Q6SPF9 oryctolagus
6	192	16.4	863	2 Q66HN3	Q66HN3 rattus norv
7	192	16.4	863	2 Q9JMD2	Q9JMD2 rattus norv
8	170	14.5	863	2 Q9JMD1	Q9JMD1 mus musculu
9	169	14.4	866	2 Q96C73	Q96C73 homo sapien
10	167	14.3	866	2 Q9HKJ3	Q9HKJ3 homo sapien
11	166.5	14.2	904	2 Q9HCF5	Q9HCF5 homo sapien
12	158	13.5	408	2 Q8CC08	Q8CC08 mus musculu
13	158	13.5	408	2 Q8OVG1	Q8OVG1 mus musculu
14	157.5	13.5	675	2 Q6DJ53	Q6DJ53 xenopus tro
15	154.5	13.2	675	2 Q6PAW6	Q6PAW6 xenopus lae
16	153.5	13.1	356	2 Q8N228	Q8N228 homo sapien
17	153	13.1	610	2 Q6N083	Q6N083 homo sapien
18	153	13.1	957	2 Q6GMQ3	Q6GMQ3 homo sapien
19	153	13.1	1004	1 PHC1 HUMAN	P78364 homo sapien
20	151	12.9	1010	2 Q7TT35	Q7TT35 mus musculu
21	151	12.9	1012	1 PHC1 MOUSE	Q64028 mus musculu
22	150	12.8	310	2 Q6TGF0	Q6TGF0 brachydanio
23	149.5	12.8	446	2 Q76866	Q76866 drosophila
24	149.5	12.8	446	2 Q9W4W7	Q9W4W7 drosophila
25	148	12.6	480	2 Q8TB59	Q8TB59 homo sapien
26	148	12.6	496	2 Q8N195	Q8N195 homo sapien
27	148	12.6	529	2 Q96NU1	Q96NU1 homo sapien
28	147	12.6	310	2 Q7ZYX8	Q7ZYX8 brachydanio
29	147	12.6	405	2 Q8BNM8	Q8BNM8 mus musculu
30	147	12.6	761	2 Q6ZPI2	Q6ZPI2 mus musculu
31	147	12.6	858	2 Q64117	Q64117 mus musculu

32	147	12.6	883	2 Q8BLB7	Q8BLB7 mus musculu
33	146.5	12.5	156	2 Q96BL4	Q96BL4 homo sapien
34	146.5	12.5	323	2 Q8N306	Q8N306 homo sapien
35	146.5	12.5	323	2 Q8K5D9	Q8K5D9 mus musculu
36	146.5	12.5	405	2 Q6NUJ6	Q6NUJ6 homo sapien
37	146.5	12.5	432	2 Q88463	Q88463 mus musculu
38	146.5	12.5	850	2 Q9QWH1	Q9QWH1 mus musculu
39	146.5	12.5	858	2 Q8IXK0	Q8IXK0 homo sapien
40	146.5	12.5	1343	2 Q06635	Q06635 bovine herp
41	146	12.5	310	2 Q8QHL4	Q8QHL4 brachydanio
42	146	12.5	827	2 Q8QHL5	Q8QHL5 brachydanio
43	143.5	12.3	623	2 Q8NA19	Q8NA19 homo sapien
44	143.5	12.3	953	2 Q7QOR1	Q7QOR1 anopheles g
45	142.5	12.2	577	2 Q9UKM5	Q9UKM5 homo sapien

ALIGNMENTS

RESULT 1					
Q96IM4					
ID	Q96IM4	PRELIMINARY;	PRT;	236 AA.	
AC	Q96IM4;				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)				
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)				
DE	LOC90378 protein (Hypothetical protein) (Fragment).				
GN	Name=LOC90378;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Villalon D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA	Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RL	and mouse cDNA sequences."				
RN	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RA	Strausberg R.;				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RA	Director MGC Project;				
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC007384; AAH07384.2; -				
DR	EMBL; BC080588; AAH0588.1; -				
DR	HSSP; P39769; 1KW4.				
DR	InterPro; IPR001660; SAM.				
DR	InterPro; IPR010993; SAM_homology.				
DR	Pfam; PF00536; SAM 1; 1.				
DR	SMART; SM00454; SAM; 1.				

[illegible]

RESULT 2
Q6POR3
ID Q6P0R3 PRELIMINARY; PRT; 285 AA.
AC Q6P0R3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypochemical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN . SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Blatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Stapleton M.J., Usdin T.B., Toshiyuki S., Carninci P., Strange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hulcy S.W.,
RA Villaion D.K., Muzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.-J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN . SEQUENCE FROM N.A.
RP RP TISSUE=Ovary;
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the ENSEMBL/GenBank/DBSJ databases.
RL EMBL: BC065477; AAH65477.1; -.
DR InterPro: IPR001660; SAM.
DR InterPro: IPR010993; SAM homologous.
DR Pfam: PF00536; SAM 1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS0105; SAM DOMAIN; 1.

KW	Hypothetical protein.
FT	NON TER 1
SQ	SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;
	Query Match 100.0%; Score 1170; DB 2; Length 285;
	Best Local Similarity 100.0%; Pred. No. 8.6e-63;
	Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 EERVLEKEEEDDEDEDVSEGSEVPESDRPAGAHQHQLNGERGPGSAKERVKEWT 60
DB	69 EERVLEKEEEDDEDEDVSEGSEVPESDRPAGAHQHQLNGERGPGSAKERVKEWT 128
QY	61 PCGPHQGQDEGRGPAPSGTTRQVFSSMAAMNKEGGTASVATGPDSPSPVLPPGKPALPGA 120
DB	129 PCGPHQGQDEGRGPAPSGTTRQVFSSMAAMNKEGGTASVATGPDSPSPVLPPGKPALPGA 188
QY	121 DGTFFGCPGKRKEPSDPEVTVMVDVVEYFTFAEGFPQOATAFOQEIDGKSLLLMQRTDV 180
DB	189 DGTFFGCPGKRKEPSDPEVTVMVDVVEYFTFAEGFPQOATAFOQEIDGKSLLLMQRTDV 248
QY	181 LTGLSIRLGPALKIYEHHIKVLQQCHPEDDDPDGFLG 217
DB	249 LTGLSIRLGPALKIYEHHIKVLQQCHPEDDDPDGFLG 285
RESULT 3	
G6PI57	PRELIMINARY; PRT; 295 AA.
ID	G6PI57
AC	G6PI57;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	LOC903078 protein (Fragment).
GN	Name=LOC903078;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
EN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Stapleton M.J., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Alschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.,
RA	Diatchenko L., Marasina K., Farmer A.A., Rubin G.W., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahey J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska I., Smalls D.B., Schmerch A., Schein J.E.,
RA	Jones S.J., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RA	Trausberg R.;
RL	Submitted (May-2002) to the ENBL/GenBank/DBJ databases.
DR	EMBL; BC030129; AAH30129.1; --
DR	InterPro; IPRO01660; SAM.
DR	Interpro; IPRO10993; SAM_homology.
DR	Pfam; PF00536; SAM 1; 1.
DR	SMART; SM00454; SAM; 1.
DR	PROSITE; PS50105; SAM DOMAIN; 1.

```

FT NON TER 1 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;
  Query Match 100.0%; Score 1170; DB 2; Length 295;
  • Best Local Similarity 100.0%; Pred. No. 8.9e-63;
  Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVKEWT 60
DB 79 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVKEWT 138

QY 61 PCGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPALPGA 120
DB 139 PCGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPALPGA 198

QY 121 DGTFFGCGPGRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMORTDV 180
DB 199 DGTFFGCGPGRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMORTDV 258

QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 259 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 295

RESULT 4
Q6SPFO PRELIMINARY; PRT; 538 AA.
AC Q6SPFO;
AT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

  Query Match 100.0%; Score 1170; DB 2; Length 538;
  Best Local Similarity 100.0%; Pred. No. 1.6e-62;
  Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVKEWT 60
DB 322 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVKEWT 381

QY 61 PCGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPALPGA 120
DB 382 PCGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPALPGA 441

QY 121 DGTFFGCGPGRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMORTDV 180
DB 442 DGTFFGCGPGRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMORTDV 501

QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 5
Q6SPF9 PRELIMINARY; PRT; 550 AA.
ID Q6SPF9
AC Q6SPF9;
AT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

  Query Match 100.0%; Score 1170; DB 2; Length 538;
  Best Local Similarity 100.0%; Pred. No. 1.6e-62;
  Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVKEWT 60
DB 322 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVKEWT 381

QY 61 PCGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPALPGA 120
DB 382 PCGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPALPGA 441

QY 121 DGTFFGCGPGRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMORTDV 180
DB 442 DGTFFGCGPGRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMORTDV 501

QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 5
Q6SPF9 PRELIMINARY; PRT; 550 AA.
ID Q6SPF9
AC Q6SPF9;
AT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

  Query Match 89.5%; Score 1047.5; DB 2; Length 550;
  Best Local Similarity 88.6%; Pred. No. 3.6e-55;
  Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLN-GERGPOSAKERVK 57
DB 331 EERVLEKEEEDDEDEDEDDVSEGEVPESDRPAQAQHHQLNGERGPOSAKERVK 390

QY 58 EWTGPHQODGRCGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPAL 117
DB 391 EWSLCPGPHGEGRGPAAGSGTRQVFSMAALSKEGSSASTTGPDSPPSPVPLPPGKPAL 450

QY 118 PGADGTPFCGPKRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMOR 177
DB 451 PGADGTPFCGPKRKEKPSDPVWTVMDVVVEYFTAGFPEQATAFQEQEIDGKSLLLMOR 510

QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 511 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 550

RESULT 6
Q66HN3 PRELIMINARY; PRT; 863 AA.
ID Q66HN3
AC Q66HN3;
AT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Sfm1 protein.
GN Name=Sfm1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC TISSUE=Testis;
RA Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Maman A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

```



```

QY 125 FGCPPGRKE-----KPSDPVEMTMDVVEYFTAGFPQATAFQOEIDGKSLLL 174
Db 767 ENKPPSPKEIRIEVDRELHDSNPLKWSADVVRFIRSTDCAPLARIFLDQEIQDQALL 826

QY 175 MORTDVLTLGSLRGLPALKIYEHKIVLQOQHFE 208
Db 827 LTLPTVQECMDLKLGAIPKLC-HHIERIKFAFYE 859

RESULT 9
Q96C73 PRELIMINARY; PRT; 866 AA.
AC Q96C73;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Scm-like with four mbt domains 1.
GN Name=SFMBT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Collins F.S., Wagner L.H., Grouse L.H., Derge J.G.,
RA Klausner R.D., Fellings P.S., Buetow K.H., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Viallan D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014614; AAH14614.1; -.
DR HSSP; P39769; 1KW4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPRO04092; MBT.
DR InterPro; IPRO01660; SAM_2.
DR Pfam; PF02820; MBT; 4.
DR SMART; SM00561; MBT; 4.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00454; SAM; 1.
SQ SEQUENCE 866 AA; 98141 MW; DCE67BF35C413BE7 CRC64;

```

```

Query Match 14.4%; Score 169; DB 2; Length 866;
Best Local Similarity 26.2%; Pred. No. 0.027;
Matches 53; Conservative 32; Mismatches 75; Indels 42; Gaps 6;

QY 10 EDDDEDEDEDDVSGSEVPESDRPAGAHQHLNGLRGFPQSAKERVKEWTPCGPHQGD 69
Db 700 EDDDDPDEGDDSLSGSTSEQD-----ELOEE-----SEMSEKSCSSSPTQS 744

```

```

QY 70 EGRGAPAGSGTRQVFSMAAMNKEGGTASV---TGPDSFSPVPLPFGKALPGADGTPFG 126
Db 745 EISTSLPPDRQR-----KELRTFSFDDENKPPSPKEIRIEVAE----- 785

QY 127 CPPGRKEKSPDPVEMTMDVVEYFTAGFPQATAFQOEIDGKSLLLMORTDVLTLGLSI 186
Db 786 ----RLHDSNPLKWSADVVRFIRSTDCAPLARIFLDQEIQDQALLLTLPTVQECMDL 841

QY 187 RLGPALKIYEHKIVLQOQHFE 208
Db 842 KLGAIPKLC-HHIERIKFAFYE 862

RESULT 10
Q9UHJ3 PRELIMINARY; PRT; 866 AA.
AC Q9UHJ3;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE RU1.
GN Name=RU1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20125026; PubMed=10661410; DOI=10.1016/S1074-7613(00)80163-6;
RA Morel S., Levy F., Bulet-Schiltz O., Brasseur F., Probst-Kepper M.,
RA Peitrequin A.L., Monsarrat B., Van Velthoven R., Cerottini J.C.,
RA Boon T., Gairin J.E., Van den Eynde B.J.;
RT "Processing of some antigens by the standard proteasome but not by the
RT immunoproteasome results in poor presentation by dendritic cells.";
RL Immunity 12:107-117(2000).
DR EMBL; AF168132; AAF19794.1; -.
DR HSSP; P39769; 1KW4.
DR Genew; HGNC:20255; SFMBT1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPRO04092; MBT.
DR InterPro; IPRO01660; SAM.
DR InterPro; IPRO11510; SAM_2.
DR Pfam; PF02820; MBT; 4.
DR Pfam; PF07647; SAM 2; 1.
DR SMART; SM00561; MBT; 4.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00454; SAM; 1.
SQ SEQUENCE 866 AA; 98229 MW; C65095C458567FC3 CRC64;

Query Match 14.3%; Score 167; DB 2; Length 866;
Best Local Similarity 26.2%; Pred. No. 0.035;
Matches 53; Conservative 31; Mismatches 76; Indels 42; Gaps 6;

QY 10 EDDDEDEDEDDVSGSEVPESDRPAGAHQHLNGLRGFPQSAKERVKEWTPCGPHQGD 69
Db 700 EDDDDPDEGDDSLSGSTSEQD-----ELOEE-----SEMSEKSCSSSPTQS 744

QY 70 EGRGAPAGSGTRQVFSMAAMNKEGGTASV---ATGPDSPVPLPFGKALPGADGTPFG 126
Db 745 EISTSLPPDRQR-----KELRTFSFDDENKPPSPKEIRIEVAE----- 785

QY 127 CPPGRKEKSPDPVEMTMDVVEYFTAGFPQATAFQOEIDGKSLLLMORTDVLTLGLSI 186
Db 786 ----RLHDSNPLKWSADVVRFIRSTDCAPLARIFLDQEIQDQALLLTLPTVQECMDL 841

QY 187 RLGPALKIYEHKIVLQOQHFE 208
Db 842 KLGAIPKLC-HHIERIKFAFYE 862

RESULT 11
Q9HCF5 PRELIMINARY; PRT; 904 AA.
ID Q9HCF5

```

```
AC O9HCF5;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE KIAA1617 protein (Fragment).
GN Name=KIAA1617;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046837; BAB13443.1; -.
DR HSSP; Q9UQR0; 1011.
DR Genew; HGNC:120256; SFMBT2.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR004092; Mbt.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR011510; SAM.2.
DR Pfam; PF02820; Mbt; 4.
DR Pfam; PF07647; SAM.2; 1.
DR SMART; SM00561; Mbt; 4.
DR SMART; SM00454; SAM; 1.
FT NON TER 1 1
SQ SEQUENCE 904 AA; 101666 MW; DB6AB35E4C58E46 CRC64;

Query Match 14.2%; Score 166.5; DB 2; Length 904;
Best Local Similarity 28.9%; Pred. No. 0.04;
Matches 58; Conservative 29; Mismatches 71; Indels 43; Gaps 8;

QY 8 EBEEDDEDEDEDDYSE--GSEVPSPDRPAGQHQLNGERGQPSQAKERVKWTGCGPH 65
DB 727 ESEESDADAMDDTASETGSEL--RDPQDTTSSAEVPSARPRRAVTLR----- 774
QY 66 QGQDEGRGAPGSGTQVFSSMAAMKEGGTASVATGPDSPSPVLPKGPALPGADGTFP 125
DB 775 SGSEPVRRPPPTTRGRGAPASSAEGE-----KCPPTX--EGTDT-- 817
QY 126 GCPGPKKEK-----PSDPVETVMQDVYFTAGFPQEQATAPQEQIDGKSLLLMQRTD 179
DB 818 -----KQEEERLVLESNPLEWTVTDVRFIKLTDCAPLAKIFQEQIDGQALLLTLP 872
QY 180 VLTGLSIRLGPALKIYEHK 200
DB 873 VQECMELKLGPAIKLC-HQIE 892

RESULT 12
ID Q8CC08 PRELIMINARY; PRT; 408 AA.
AC Q8CC08;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
DE library, clone:9330161D17 product:hypothetical SAM domain (Sterile
DE alpha motif) containing protein, full insert sequence.
GN Name=Scml4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
```

```
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Hashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (KISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK034171; BAC28615.1; -.
DR HSSP; P39769; 1KW4.
DR MGD; MGI:2446140; Scml4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000637; A+T_hook.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM.1; 1.
DR SMART; SM00384; A+T_hook; 1.
DR SMART; SM00454; SAM; 1.
KW Hypothetical protein.
SQ SEQUENCE 408 AA; 44445 MW; 59DD91591D10D59 CRC64;

Query Match 13.5%; Score 158; DB 2; Length 408;
Best Local Similarity 26.9%; Pred. No. 0.059;
```

Qy	14	DEBDEDDVSEGSFESP----	RAPCAQHQLN--GRCPSQAKERVKWETPCGHOGQ	68
Dd	215	DKAQREDGGTESAKVATAECLANAVGMNRyamDFSHRGSvVTHSSSLYKRLTCG---	DS	271
Qy	69	DEGRGPA-POSGTRQVFSAAMNKEGTASVATGPDPSPVPPLPPGPKPALPGADGTPFGC	127	
Dd	272	HLAGPATTTSGR-----	TNPVPSGGSSPGLRLPASSPKNGTAIEGNRC	318
Qy	128	PPG-----	RKEKPSDPEVTMDVVEYFTAGFPE----QAATAFQEIDGKSILLM	175
Dd	319	APSPPSPEVDQTRPRSSRNPTMTVEDVVRVKDAD--PEALGPHVELFRKHIEDIGNALLL	377	
Qy	176	QRTDVLTLGISIRLGPKALKIYEHHIKVLQQGFH	207	
Dd	378	RSDMIKMYYLKLGLGPKALKLC-YHIDILKQAKF	408	

RESULT 14
Q6DJ53 PRELIMINARY; PRT; 675 AA.

ID	Q6DJ53	AC	Q6DJ53
DT	25-OCT-2004	(T-EMBLrel. 28, Created)	
DT	25-OCT-2004	(T-EMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(T-EMBLrel. 28, Last annotation update)	
DE	Scmhl-prov protein.		
DN	Name=scmhl-prov;		
OS	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCHI_TaxID=8364;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=Whole body;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Scheraga H.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.P., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Topolycki S., Carninci P., Prange C., Raja S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bogak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Musny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.; Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schultz J., Myers R.M., Butterfield Y.S., Kravinsky M.I., Skalska U., Smallos D.E., Schnier A., Schein J.E., Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=Whole body;		
RC	Klein S., Gerhard D.S.;		
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.		
RR	EMBL; BC075330; AAA75330.1;		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0045449; P:regulation of transcription; IEA.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR004092; MbT.		
DR	InterPro; IPR001660; SAM.		
DR	InterPro; IPR011510; SAM 2.		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 24.4357 Seconds
(without alignments)
854.447 Million cell updates/sec

Title: US-10-671-242-7

Perfect score: 1170

Sequence: 1 BERVLSKBEEDDEDEDEE.....HIKVLQGGHFEDDDPGFLG 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	12.9	1012	2 I53172	RAE-28 - mouse
2	136	11.6	1300	2 T03166	probable immediate
3	135	11.5	1589	2 T13506	hypothetical prote
4	130	11.1	930	2 D37271	A-alpha 2 4 protei
5	124.5	10.6	2453	2 S60254	nuclear receptor c
6	123.5	10.6	783	2 A55817	cyclin-dependent k
7	123	10.5	300	2 S19560	proline-rich prote
8	123	10.5	754	2 A56619	female sterile hom
9	122.5	10.5	206	1 P1RT3	acidic proline-ric
10	122.5	10.5	295	2 B48013	proline-rich prote
11	121.5	10.4	768	2 H54024	protein kinase (EC
12	121.5	10.4	777	2 F54024	protein kinase (EC
13	121.5	10.4	777	2 B54024	protein kinase (EC
14	121.5	10.4	779	2 E54024	protein kinase (EC
15	119.5	10.2	367	1 OZZQMY	circumsporozoite p
16	118.5	10.1	1477	2 T13797	tumor suppressor pr
17	118	10.1	1280	2 T00365	hypothetical prote
18	117.5	10.0	1282	2 JE0120	glycoprotein A - m
19	116	9.9	301	2 E29149	proline-rich prote
20	115.5	9.9	617	2 S42719	actin-binding prot
21	115	9.8	772	2 I50463	protein kinase - c
22	115	9.8	1840	2 T30250	GPI protein - mous
23	114.5	9.8	280	2 S35103	bone sialoprotein
24	114.5	9.8	1460	1 EDBBIF	immediate-early pr
25	114	9.7	909	2 S32538	cGMP-gated cation
26	114	9.7	1388	2 A53317	collagen alpha 1(X
27	113	9.7	226	2 S41032	hypothetical prote
28	113	9.7	916	2 A27864	neurofilament trip
29	112.5	9.6	512	2 E59437	F02569_2 protein {

30 112.5 9.6 647 2 T43952
31 112 9.6 786 2 A35466
32 112 9.6 892 2 B46203
33 111.5 9.5 1494 2 T14355
34 111.5 9.5 1880 2 T18531
35 111 9.5 1252 2 T14272
36 110 9.4 1110 2 I51116
37 110 9.4 1213 2 S16356
38 109.5 9.4 730 2 A36226
39 109 9.3 902 2 T26775
40 108 9.2 1770 2 A48013
41 107.5 9.2 1315 2 A56101
42 107.5 9.2 1774 2 B56101
43 107 9.1 907 2 A45560
44 107 9.1 2924 2 T18378
45 106.5 9.1 299 2 A29681

hypothetical prote
progesterone recep
mating type A alph
protein-tyrosine-p
tractin - medicina
cortactin-binding
NF-180 - sea lamp
ovo protein - frui
collagen alpha 1 c
hypothetical prote
proline-rich prote
collagen alpha 1(X
collagen alpha 1(X
sporozoite surface
variant-specific s
nucleophosmin - Af

ALIGNMENTS

RESULT 1

I53172

RAE-28 - mouse

C;Species: Mus sp. (mouse)

C;Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 23-Sep-2002

C;Accession: I53172; I66850; I66851

R;Nomura, M.; Takihara, Y.; Shimada, K.

Differentiation 57, 39-50, 1994

A;Title: Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cell

Drosophila polyhomeotic protein.

A;Reference number: I53172; MUID:94350162; PMID:8070621

A;Accession: I53172

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1012 <RES>

A;Cross-references: GB:S73882; NID:G688186; PIDN:AAB31766.1; PID:G688187

A;Accession: I66850

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-75, 'V' <RE2>

A;Cross-references: GB:S73883; NID:G688188

A;Accession: I66851

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-75, 'V' <RE2>

A;Cross-references: GB:S73884; NID:G688189

C;Genetics:

A;Gene: rae-28

C;Superfamily: LAR-interacting protein; SAM homology

F;373-476/Region: Glutamine-rich

F;945-1011/Domain: SAM homology <SAM>

Query Match 12.9%; Score 151; DB 2; Length 1012;

Best Local Similarity 27.8%; Pred. NO. 0.0063;

Matches 49; Conservative 27; Mismatches 54; Indels 46; Gaps 8;

QY 47 RGPQSAKERVKWTGCPG-HQCG-DEGRGPAQSGTRQVPSMAAMNKEGKTASVATGPD 104

DB 864 RGPRESSSIIARAKTQGRHGOESSRG-----SDNSSYDEALSPTS 906

QY 105 PGFVPLPPG-----KPALPGADGTFPGCPGRKEKP-----SDPVTWMDVVEY 149

DB 907 PGFLSVRAGHGRLDGLNTTTPSTPELQGI-----NPVFLSNPSQSWSEVVEYEF 956

QY 150 FTE-AGFPQATAPQEQEIDGKSLLMQRTDVLTLGSLRGLPALKTYHHIKVLQ 204

DB 957 IASLQGCQBIAEFRSQEIDGQALLLKEHLMSAMNKLGPALKICA-KINVLKE 1011

RESULT 2

T03166

probable immediate early protein - alcelaphine herpesvirus 1

C:Species: alcelaphine herpesvirus 1
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03166
R:Ensser, A.; Pflanz, R.; Fleckenstein, B.
J. Virol. 71, 6517-6525, 1997
A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.
A:Reference number: Z14840; MUID:97404659; PMID:9261371
A:Accession: T03166
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1300 <ENS>
A:Cross-references: UNIPROT:O36421; EMBL:AF005370; NID:G2337967; PIDN:AA058118.1; PID:G2337967
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 11.6%; Score 136; DB 2; Length 1300;
Best Local Similarity 31.2%; Pred. No. 0.089;
Matches 50; Conservative 10; Mismatches 66; Indels 34; Gaps 8;

QY 8 EEEEDDEDEDDVSEG---SEVPESDRPAGAAHQHQLNGE--RGQSAKERVKEWTPC 62
Db 615 EGDGEPPEEGPGRGSDPGAGQEVPEG--FKGFE-----GECQSPSSCEGQQVPKGPD 667
QY 63 GPHQ-----QODRGGPA-PGSGTRQVFSMAAMKEGTTASVATGDPDPSPLPPGK-- 114
Db 668 GPEGSSGGSSGEGSPGSGEGQGVKAGSGEGPCRPGGDEDDGDPGPDTEG 727
QY 115 -----PALPGADGTPFG-----CPPGRKEKPSDPVE 140
Db 728 EGPCGPGPDGDPGPDGTEGEGPCGPGPDGDPGEE 767

RESULT 3
T13606
hypothetical protein 87B1.5 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13606; S23632
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17668
A:Accession: T13606
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1589 <MUR>
A:Cross-references: UNIPROT:P39769; EMBL:Z98269; NID:el355202; PID:el251078; PIDN:CAB109
R:DeCamillis, M.; Cheng, N.; Pierre, D.; Brock, H.W.
Genes Dev. 6, 223-232, 1992
A:Title: The polyhomeotic gene of Drosophila encodes a chromatin protein that shares pol
A:Reference number: S23632; MUID:92146957; PMID:1346609
A:Accession: S23632
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1010,'V',1012-1192,'L',1194-1274,'I',1276-1589 <DEC>
A:Cross-references: EMBL:X63672; NID:g11056; PIDN:CAA45211.1; PID:g11057
C:Genetics:
A:Gene: FlyBase:ph-p
A:Cross-references: FlyBase:FBgn0004861; FlyBase:FBgn0004860
A:Introns: 12/2; 595/1; 745/2; 1340/1
C:Superfamily: LAR-interacting protein; SAM homology
C:Keywords: DNA binding; nucleus
F:74-80,247-285,411-450,494-650,727-737,775-955,1032-1061/Region: glutamine-rich
F:1510-1576/Domain: SAM homology <SAM>

Query Match 11.5%; Score 135; DB 2; Length 1589;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 60; Conservative 38; Mismatches 86; Indels 66; Gaps 12;

QY 2 ERVLEKEEEDDEDE---DEEDDVSEGVESVPSDRPAGAAHQHQLNGERGQPSAKERVKE 58
Db 1327 QRYADKDVSDPEPPKKATWQEDIKLSGIASAFGSDMVACEQ-----CGKMEHKAALKR 1379
QY 59 WTPCGP---HQGD---EGRGPAPGSGTRQVFSMAAM-----NKEGGT----- 95

C:Species: Schizophyllum commune
C:Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 09-Jul-2004
C:Accession: D37271
R:Ullrich, R.C.
submitted to the Protein Sequence Database, October 1991
A:Reference number: A37271
A:Accession: D37271
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-930 <ULL>
A:Cross-references: UNIPROT:P37938
C:Keywords: DNA binding; nucleus; transcription regulation

Query Match 11.1%; Score 130; DB 2; Length 930;
Best Local Similarity 22.4%; Pred. No. 0.16;
Matches 55; Conservative 23; Mismatches 84; Indels 84; Gaps 9;

QY 1 ERVLEKEEEDDEDEDDVSEGVESVPSDRPAGAAHQHQLNGER-GPQSAKERVKEW 59
Db 429 DESTDEDDSDSENDSDEDEDEDEDEEEEPVKIAGAKRGNDDEEVSFLAKPRIFS- 487
QY 60 TPCGPHQGDGEGPAGSGTRQVFSMAAMKEGTTASVATGDPDPSPV----- 108
Db 488 PVRPRPQAIRVSLFSPAPSSR-----GSTPTSPVSPSPKAKRPAQATSL 533
QY 109 ---PL-----PPKPALPGADGTPF----- 125
Db 534 ASHPMKREKLEELRKAGLAPPSAPVLMGPDGVLGTVRSRSPSPSPSVSLPLP 593
QY 126 --GCPPGRKEKPSDPVEVTVMVDVVEYFTAGPFEQATAFQEQE-----IDGKSL 174
Db 594 SRGVSGGKVTGDTFPVWNWD-LEAHTQA--PRDLTAATKSSAGCSVDVPLPGKSRSL 650
QY 175 MQRTDV 180
Db 651 TRSPSI 656

RESULT 5
S60254
nuclear receptor co-repressor N-COR - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S60254
R:Hoerlein, A.J.; Naeae, A.M.; Heinzel, T.; Torchia, J.; Gloss, B.; Kurokawa, R.; Ryan,
Nature 377, 397-404, 1995
A:Title: Ligand-independent repression by the thyroid hormone receptor mediated by a nuc
A:Reference number: S60254; MUID:96008539; PMID:7566114
A:Accession: S60254
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2453 <HOB>
A:Cross-references: UNIPROT:Q60974; EMBL:U35312; NID:g1022717; PIDN:AA017125.1; PID:g1022

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <ROB>
A:CROSS-references: UNIPROT:Q61888; GB:X58438; NID:g53181; PIDN:CAA41344.1; PID:g53182
R:Roberts, S.G.E.; Layfield, R.; McDonald, C.J.
Nucleic Acids Res. 19, 5205-5211, 1991
A>Title: The mouse prolamine-rich protein MP6 promoter binds isoprenaline-inducible parotid
A:Reference number: S22570; MUID:92020206; PMID:1747160
A:Accession: S22570
A:Molecule type: DNA
A:Residues: 1-14 <RO2>
A:CROSS-references: EMBL:X61126
C:Superfamily: proline-rich protein

Query Match 10.5%; Score 123; DB 2; Length 300;
Best Local Similarity 29.5%; Pred. No. 0.15;
Matches 43; Conservative 10; Mismatches 61; Indels 32; Gaps 7;

Qy 8 BEEDDDDEDEDDVSEGSVPESDRPAGAHH---QLNGERGPGSAKRVKWTTCGP 64
 :|:|||||:|||||:
Db 62 DENNDDEDGSEED--VNREPGPPHPPHSGNNHAPPQQGDAGHPRPNQGGPPSPGPP 119
 :|:|||||:|||||:
Qy 65 HQQDEGRGAPGCTGTQVFSAANKEGTSATVGPDSPSP-----VPLPPG----KPA 116
 :|:|||||:|||||:
Db 120 PQSSSQRRPPOFG-----NQQPPPGQGQQRPPQGNQGGPPPPGGFQQRPP 167
 :|:|||||:|||||:
Qy 117 LPGAD--GTPFGCP-----PGRKEKP 135
 :|:|||||:|||||:
Db 168 QPGNGGGPPQGGPHPPRPFGNQOGP 193
 :|:|||||:|||||:

RESULT 8
A56619
female sterile homeotic (fish) homolog RING3 - human
C:Species: Homo sapiens (man)
C>Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 20-Sep-1999
C:Accession: A56619; S18860; S40781
R:Beck, S.; Hanson, I.; Kelly, A.; Pappin, D.J.; Trowsdale, J.
DNA Seq. 2, 203-210, 1992
A>Title: A homologue of the Drosophila female sterile homeotic (fsh) gene in the class I
A:Reference number: A56619; MUID:92229974; PMID:1352711
A:Accession: A56619
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-754 <BEC>
A:CROSS-references: EMBL:X62083; NID:g31471; PIDN:AAA68930.1; PID:g31472; EMBL:M80613; N
A>Note: sequence extracted from NCBI backbone (NCBIP:108781)
C:Genetics:
A:Gene: RING3
C:Superfamily: unassigned bromodomain proteins; bromodomain homology
C:Keywords: duplication
E:52-109/Domain: bromodomain homology <BRO1>
F:325-382/Domain: bromodomain homology <BRO2>

Query Match 10.5%; Score 123; DB 2; Length 754;
Best Local Similarity 29.2%; Pred. No. 0.39;
Matches 47; Conservative 22; Mismatches 58; Indels 34; Gaps 8;

Qy 7 KEEEDDDDEDEDDVSEGSVPESDRPAGAAQ-----HHQLNG-ERGPQS-----AK 53
 :|:|||||:|||||:
Db 445 EEEDDEDEDEBESESSESSEEAHRLAELQEQLRAVHEQLAALSQGPISKPKRKREK 504
 :|:|||||:|||||:
Qy 54 ERVKWPTPCGHQCG-----DEGRGAPAGTGTRQVFSAAMNKEGTSATVAT---GPDSPS 106
 :|:|||||:|||||:
Db 505 KEKKKKRKAERKHARGAGAEDDDKGPRAPRPPOPKSKKASGSGGGAALGPSGFPSGGGS 564
 :|:|||||:|||||:
Qy 107 VPVLPP-----PGKPALP-GADGTGFCPPGRKEKPSDPVEW 141
 :|:|||||:|||||:
Db 565 GTKLPKXATKTAPPALPTYD-----SEEEEESRPMYSY 597
 :|:|||||:|||||:

RESULT 9
PIRT3

acidic proline-rich protein precursor - rat
N;Alternate names: PRP
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C;Accession: A03296
R;Ziemer, M.A.; Swain, W.F.; Rutter, W.J.; Clements, S.; Ann, D.K.; Carlsson, D.M.
J. Biol. Chem. 259, 10475-10480, 1984
A;Title: Nucleotide sequence analysis of a proline-rich protein cDNA and peptide homolog
A;Reference number: A03296; MUID:84289443; PMID:6547951
A;Accession: A03296
A;Molecule type: mRNA
A;Residues: 1-206 <2IE>
A;Cross-references: UNIPROT:P04474; GB:K02247; NID:g206395; PIDN:AAA41949.1; PID:g206396
C;Comment: This protein contains six 18- to 19-residue repeats.
C;Comment: This protein may protect teeth by binding to tannins.
C;Superfamily: proline-rich protein
C;Keywords: duplication; parotid gland; saliva; tandem repeat
F;1-13/Domain: signal sequence #status predicted <SIG>
F;14-206/Product: acidic proline-rich protein #status predicted <MAT>
F;80-189/Region: 18-residue repeats

Query Match 10.5%; Score 122.5; DB 1; Length 206;
Best Local Similarity 26.4%; Pred. No. 0.11;
Matches 39; Conservative 14; Mismatches 60; Indels 35; Gaps 6;

QY 8 EEEEDDEDEDDVSEGEVPE-----SDRPAGAHQHQLNGERGPOSASAKERVKEW 59
DB 55 ENGDDGDDSDGDDGDDGNRPERRPPHGGNHNQRPDPGHHGPPSGGPGQTSQSGNPQ 114

QY 60 TPCGPHQGQDEG-----RGPAFGSGTRQVFSMAAMNKEGTSATVATGDSPPVPLPP 112
DB 115 GP--PFGGPGQGPFGPNPGPPP-----QGGFQQRPPQGGKPGGP-PQ 156

QY 113 GKPALPGADGTPFGCPP--GRKEKPSDP 138
DB 157 GGGFPPPGPNPGPPPPGQGHQQRPPQ 184

RESULT 10
B48013
proline-rich proteoglycan 2 precursor, parotid - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 09-Jul-2004
C;Accession: B48013
R;Castle, A.M.; Castle, J.D.
J. Biol. Chem. 268, 20490-20496, 1993
A;Title: Novel secretory proline-rich proteoglycans from rat parotid. Cloning and charac
A;Reference number: A48013; MUID:93388626; PMID:8376404
A;Accession: B48013
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-295 <CAS>
A;Cross-references: UNIPROT:Q07611; GB:L17318; NID:g310199; PIDN:AAA03074.1; PID:g310200
C;Superfamily: proline-rich protein
C;Keywords: extracellular protein; glycoprotein; tandem repeat

Query Match 10.5%; Score 122.5; DB 2; Length 295;
Best Local Similarity 26.1%; Pred. No. 0.16;
Matches 43; Conservative 15; Mismatches 68; Indels 39; Gaps 7;

QY 9 EEEEDDEDEDDVSEGSVPESDRPA--GAQHQQ-----LNGERGPOSASAKERVKEW 59
DB 61 DENGDDNDGDDGDDGDDNRPERRPPQHGGNHHHPHPPAAGPQRPQPGSPQGGPP- 119

QY 60 TPCGPHQGQDEG-----RGPAFGSGTRQVFSMAAMNKEGTSATVATG-P 102
DB 120 PPGGPQRRPPQGGPQRPQPGSPQGPQPPPGGQRPQPPPGPQRPQPGSP 179

QY 103 DSPSPV-----PLPGKALPGADGTPFGCPP--GRKEKP 135
DB 180 QGPPPPGGPQQRAPQGPQPPPGGQRRPQPGSPQGPQPPPGPQQRP 224

RESULT 11
H54024
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE alpha 2-3 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Aug-2004
C;Accession: H54024; GS4024
R;Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.
J. Biol. Chem. 269, 15786-15794, 1994
A;Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinase
A;Reference number: A54024; MUID:94253170; PMID:8195233
A;Accession: H54024
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-768 <XIA>
A;Cross-references: GB:U04817
A;Note: authors translated the codon AAA for residue 5 as Leu, GAC for residue 6 as Lys,
CTC for residue 342 as Phe, CTG for residue 574 as Gln, AAG for residue 614 as Asn, GAT 1
A;Accession: G54024
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-114,117-768 <X12>
A;Cross-references: GB:U07705
A;Note: authors translated the codon AAG for residue 5 as Leu, GAC for residue 6 as Lys,
CTC for residue 351 as Phe, GAT for residue 418 as Glu, AAC for residue 457 as Lys, GTA 1
C;Superfamily: protein kinase homology
C;Keywords: alternative splicing; ATP; phosphotransferase
F;409-664/Domain: protein kinase homology <KIN>
F;417-425/Region: protein kinase ATP-binding motif

Query Match 10.4%; Score 121.5; DB 2; Length 768;
Best Local Similarity 26.0%; Pred. No. 0.51;
Matches 63; Conservative 31; Mismatches 83; Indels 65; Gaps 13;

QY 8 EEEEDDEDEDDVSEGEVPESDRPAGAHQHQLNGERGPOSASAKERVKEWTPCGPHQ- 66
DB 280 EEEEEEEEEEGSTSESESEEEEE---EETGTSNSEASESQSAEEVSEEMSEDEER 336

QY 67 -----GDEGRGAPGSGTRQVFSMAAMNKEGTSATVATGPDSP--S 106
DB 337 ENENHFLVVPESRFRDRDSGESEAEAEVGEVGTQSSALT-----EGDYV-----PDSFALS 387

QY 107 PV-----PLPPGKPALPGA-----DGTPEGCPGRKEKPSDPVETVMDVVEY 149
DB 388 PIELKQELPKYLPALQGCSEVEEFQCLNRIEST-YGVVYRAKDKYTDRIALKRLKMEK 446

QY 150 FTEAGFPEQATAPQEQEIDGKSLLMQRTDVTGLSIRLGPAL-KIY-----EHKIKVL 202
DB 447 EKE-GFP--ITSIREI-----NTILKAQHPIVTVREIVVGSNMCKIYIVMNYVEHDLKSL 499

QY 203 QQ 204
DB 500 ME 501

RESULT 12
F54024
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE beta 2-1 - human
N;Contains: protein kinase cdc2-related PITSLRE beta-1
C;Species: Homo sapiens (man)
C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Aug-2004
C;Accession: F54024; D54024
R;Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.
J. Biol. Chem. 269, 15786-15794, 1994
A;Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinase
A;Reference number: A54024; MUID:94253170; PMID:8195233
A;Accession: F54024
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-777 <XIA>
A;Cross-references: GB:U07704
A;Accession: D54024
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA

Qy 150 FTEAGFPEQATAFOEQEIDGKSLLLMQRTDVLTLGSLRLGPAL-KIY-----EHHIKVL 202
 Db 456 EKE-GFP--ITSLSREI-----NTILKAQHPNIVTVREIVVGSNMCKIYIVMNVYVEHDLKSL 508

Qy 203 QQ 204
 Db 509 ME 510

RESULT 14
 E54024
 protein kinase (EC 2.7.1.37) cdc2-related PITSLRE alpha 2-1 - human
 C/Species: Homo sapiens (man)
 C/Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Aug-2004
 C/Accession: E54024
 R/Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.
 J. Biol. Chem. 269, 15786-15794, 1994
 A/Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinase
 A/Reference number: A54024; MUID:94253170; PMID:8195233
 A/Accession: E54024
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-779 <XIA>
 A/Cross-references: GB:U04824
 C/Superfamily: protein kinase homology
 C/Keywords: alternative splicing; ATP; phosphotransferase
 F:420-675/Domain: protein kinase homology <KIN>
 F:428-436/Region: protein kinase ATP-binding motif

Query Match 10.4%; Score 121.5; DB 2; Length 779;
 Best Local Similarity 26.0%; Pred. NO. 0.51;
 Matches 63; Conservative 31; Mismatches 83; Indels 65; Gaps 13;

Qy 8 EEREDDEDEEDVDGSEVPSDRPAGAQHQLNGERGQSAKERVKEWTPCGPHQ- 66
 Db 291 EEEEEEEEEEGSTSESEEEEE---EETGTSNSEEASEQSAEEVSEEMSEDER 347

Qy 67 -----GQDEGRGPAPGSGTRQVFSMAAMNKEGQTASVATGPDSP--S 106
 Db 348 ENENHFLVPSRFRDPSGESEAEVEEGTGPQSSALT-----EGDYV-----PDSFALS 398

Qy 107 PV-----PLPPGKPALPGA-----DGTFFGCPGGRKEKPSDPVEWTVMDVVEY 149
 Db 399 PIELKQELPKYLPALQGRSVEEFQCLNRIEETG-YGVVYRAKDKKTDEIVALKRLKWEK 457

Qy 150 FTEAGFPEQATAFOEQEIDGKSLLLMQRTDVLTLGSLRLGPAL-KIY-----EHHIKVL 202
 Db 458 EKE-GFP--ITSLSREI-----NTILKAQHPNIVTVREIVVGSNMCKIYIVMNVYVEHDLKSL 510

Qy 203 QQ 204
 Db 511 ME 512

RESULT 15
 OZZOMY
 circumsporozoite protein precursor - Plasmodium yoelii
 N/Alternate names: sporozoite surface antigen
 C/Species: Plasmodium yoelii
 C/Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C/Accession: A26271
 R/Jal, A.A.; de la Cruz, V.F.; Welsh, J.A.; Charoenvit, Y.; Maloy, W.L.; McCutchan, T.F.
 J. Biol. Chem. 262, 2937-2940, 1987
 A/Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii
 A/Reference number: A26271; MUID:87137555; PMID:3102479
 A/Accession: A26271
 A/Molecule type: DNA
 A/Residues: 1-367 <LAL>
 A/Cross-references: GB:U02695; NID:g150222; PIDN:AAA29558.1; PID:g160223
 C/Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
 C/Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repeat.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-367/Product: circumsporozoite protein #status predicted <MAT>
F;139-228/Region: 6-residue repeats (Q-G-P-G-A-P)
F;229-260/Region: 4-residue repeats (Q-Q-P-P)
F;293-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match	10.2%	Score 119.5;	DB 1;	Length 367;
Best Local Similarity	31.4%	Pred. NO. 0.32;		
Matches	44;	Conservative 13;	Mismatches 50;	Indels 33; Gaps 8;

QY	5	LEKEEEEDDEDEDDVSEGESEVPESDRPAGAQHHOLNGERGQPSAKERVKE--WTPC	62
Db	92	LPKEKKDDLPKEEKDDPPKD---PKDDPPKEAQNKLN---QPVVADENVDOGFGAPQ	145
QY	63	GPHQGQDEGRGPAPGSGTRQVFSMAANKEGGTASVATGPDSPS--FVPLPPGKPALPGA	120
Db	146	GP--GAPQGGFAPQGGFAPQ-----GFGAPQGGFAPQGGFAPQGGFAPQ	185
QY	121	DGTPFGCP--PGRKEKPSDP	138
Db	186	PQGP-GAPQGGFAPQGGFAP	204

Search completed: September 20, 2005, 12:43:18
Job time : 25.4357 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 8.92289 Seconds
(without alignments)
650.171 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82
Sequence: 1 EEEEDDEDEDD 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	15	4	AAB82809 Human low
2	82	100.0	26	4	AAB82810 Human low
3	82	100.0	217	2	Aaw49041 Human low
4	82	100.0	217	4	AAB82803 Human low
5	82	100.0	241	3	AAB88957 Breast an
6	82	100.0	538	4	AAB82806 Human low
7	75	91.5	358	6	Abb84611 Soybean n
8	73	89.0	457	3	Aay50875 Murine NN
9	73	89.0	1085	2	Aar95607 RRP3 telo
10	73	89.0	1085	7	Adk62464 Disease t
11	72	87.8	114	5	Abg93347 C. albica
12	72	87.8	178	3	Aag29185 Arabidops
13	72	87.8	178	3	Aag34230 Arabidops
14	72	87.8	178	8	Adn74341 Thale cre
15	72	87.8	178	8	Adn73879 Thale cre
16	72	87.8	288	8	Adn24238 Bacterial
17	72	87.8	427	5	Abp73549 Candida a
18	71	86.6	376	5	Adi17117 Human NOV
19	71	86.6	474	5	Adi17115 Human NOV
20	71	86.6	587	8	Adh72068 Human pro
21	71	86.6	589	8	Adr09338 Human pro
22	71	86.6	591	8	Adh72056 Human pro
23	71	86.6	591	8	Adh72060 Human pro
24	71	86.6	596	8	Adh72058 Human pro
25	71	86.6	597	7	Add89035 TAT276. 1

26	71	86.6	598	8	ADN02703	Adn02703 Human rec
27	71	86.6	612	5	AD117114	Ad117114 Human NOV
28	71	86.6	612	6	ABG74699	Abg74699 Human CGD
29	71	86.6	634	8	ADJ67009	Adj67009 Human sec
30	71	86.6	638	6	ABG74700	Abg74700 Human CGD
31	71	86.6	642	8	ADH72066	Adh72066 Human pro
32	71	86.6	646	8	ADH72062	Adh72062 Human pro
33	71	86.6	646	8	ADH72064	Adh72064 Human pro
34	71	86.6	652	6	ABG74695	Abg74695 Human CGD
35	71	86.6	667	5	AD117113	Ad117113 Human NOV
36	71	86.6	667	8	AD018006	Adq18006 Human sof
37	71	86.6	667	8	ABM81423	Abm81423 Tumour-as
38	71	86.6	668	5	AD116706	Adi16706 Human NOV
39	71	86.6	668	8	ADH72054	Adh72054 Human pro
40	71	86.6	668	8	ADH72072	Adh72072 Human pro
41	71	86.6	668	8	ADH72074	Adh72074 Human pro
42	71	86.6	668	8	ADN42360	Adn42360 Human nov
43	71	86.6	712	7	AD61623	Ade61623 Rat Prote
44	71	86.6	712	7	ADD46090	Add46090 Rat Prote
45	71	86.6	712	7	ADE57828	Ade57828 Rat Prote

ALIGNMENTS

RESULT 1
AAB82809
ID AAB82809 standard; peptide; 15 AA.
XX AAB82809;
XX
DT 12-NOV-2001 (first entry)
XX
DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.
XX
KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
XX atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX Homo sapiens.
XX
PN WO200164874-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006356.
XX
PR 02-MAR-2000; 2000US-00517849.
PR 14-JUL-2000; 2000US-00616289.
(BOST-) BOSTON HEART FOUND INC.
PI Lees AM, Lees RS, Law SW, Arjona AA;
XX
XX WPI; 2001-565505/63.
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
diagnosing and/or identifying therapeutic agents for atherosclerosis.
PS Claim 14(a); Page; 143pp; English.
XX
XX The present sequence is that of a peptide comprising amino acid residues
329-343 of novel human low density lipoprotein binding protein 2 (LBP-2,
see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of
the invention that are capable of binding to native and methylated low
density lipoproteins (LDLs). Also claimed are biologically active
fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
as expression vectors, cells and methods of producing the LBPs.
CC Polypeptides having the present amino acid sequence are among those
claimed. Methods of determining if an animal is at risk for
CC atherosclerosis, methods for evaluating an agent for use in treating
CC atherosclerosis, and methods for treating a cell having an abnormality in
CC structure or metabolism of LBP are also claimed, as are pharmaceutical
CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine

CC compositions. Note: the present sequence is not shown in the
 CC specification but is derived from the human LBP-2 sequence given in
 CC figure 7A (see AAB82806)

XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 82; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||
 DB 1 EEEEDDEDEDEDD 15

RESULT 2
 ID AAB82810
 XX AAB82810 standard; peptide; 26 AA.

AC AAB82810;
 DT 12-NOV-2001 (first entry)
 XX Human low density lipoprotein binding protein 2 (LBP-2) peptide.
 DE
 XX
 XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 XX Homo sapiens.

OS
 PN WO200164874-A2.

XX
 PD 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX
 PA (BOST-) BOSTON HEART FOUND INC.

XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;

XX
 DR WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 14(a); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues
 CC 329-354 of novel human low density lipoprotein binding protein 2 (LBP-2,
 CC see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of
 CC the invention that are capable of binding to native and methylated low
 CC density lipoproteins (LDLs). Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs.
 CC Polypeptides having the present amino acid sequence are among those
 CC claimed. Methods of determining if an animal is at risk for
 CC atherosclerosis, methods for evaluating an agent for use in treating
 CC atherosclerosis, and methods for treating a cell having an abnormality in
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
 CC compositions. Note: the present sequence is not shown in the
 CC specification but is derived from the human LBP-2 sequence given in
 CC figure 7A (see AAB82806)

XX
 SQ Sequence 26 AA;

Query Match 100.0%; Score 82; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.0026;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||
 DB 1 EEEEDDEDEDEDD 15

RESULT 3
 AAW49041

ID AAW49041 standard; protein; 217 AA.

XX AAW49041;

XX 09-NOV-1998 (first entry)

XX Human low density lipoprotein binding protein LBP-2.

XX Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
 KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 8..33 /note= "Claim 2"

FT Peptide 8..32 /note= "Claim 2"

FT Peptide 23..33 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

XX WO9823282-A1.

XX 04-JUN-1998.

XX 26-NOV-1997; 97WO-US021857.

XX 27-NOV-1996; 96US-0031930P.

PR 03-JUN-1997; 97US-0048547P.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 1998-322455/28.

DR N-PSDB; AAV32838.

XX Nucleic acid encoding low density lipoprotein binding proteins and
 PT related vectors - transformed cells, proteins, and modulators of binding,
 PT useful for treatment and diagnosis of atherosclerosis and for identifying
 PT subjects at risk.

XX Claim 1; Fig 7; 47pp; English.

XX This polypeptide comprises novel human low density lipoprotein (LDL)
 CC binding protein LBP-2 that is capable of binding both native and methyl
 CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
 CC AAV32838). CDNA clones (see AAV32834-39) and encoded rabbit and human
 CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP
 CC invention provides methods for determining if an animal is at risk for
 CC atherosclerosis (e.g. for prenatal screening); methods for treating
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
 CC methods for treating a cell having an abnormality in LBP structure or
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
 CC well as recombinant vectors and host cells used to produce recombinant
 CC LBP

XX Sequence 217 AA;

Query Match 100.0%; Score 82; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D1	XX	Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.
DE	XX	
KW	XX	Human, breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW	XX	neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW	XX	antidiabetic; antiinflammatory; antitumor; vulnary; anticoagulant;
KW	XX	antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW	XX	Addison's disease; allergy; autoimmune haemolytic anaemia;
KW	XX	autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW	XX	multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW	XX	cardiovascular disorder; wound healing; neurological disease.
OS	XX	Homo sapiens.
XX	XX	
FN	XX	WO2000055173-A1.
XX	XX	
PD	XX	21-SEP-2000.
XX	XX	
PF	XX	08-MAR-2000; 2000WO-US005881.
XX	XX	
PR	XX	12-MAR-1999; 99US-0124270P.
XX	XX	
PA	XX	(HUMA-) HUMAN GENOME SCI INC.
XX	XX	
PI	XX	Rosen CA, Ruben SM;
XX	XX	
DR	XX	WPI, 2000-611515/58.
N-	XX	PSDB; AAF21860.
XX	XX	
PT	XX	New human breast and ovarian cancer associated gene sequences and the
PT	XX	polypeptides encoded by these genes, useful in the prevention, treatment
FT	XX	and diagnosis of cancer, immune disorders, cardiovascular disorders and
PT	XX	neurological diseases.
XX	XX	
PS	XX	Claim 11; Page 1112; 1299pp; English.
XX	XX	
CC	XX	Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC	XX	proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC	XX	associated with breast and ovarian cancer. Included in the invention are
CC	XX	sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC	XX	isolation and characterisation of the DNA and protein sequences of the
CC	XX	invention. The breast and ovarian cancer associated DNA, protein, agonist
CC	XX	or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic;
CC	XX	neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC	XX	antiinflammatory; antitumor; vulnary; anticoagulant; antibacterial;
CC	XX	antifungal; antiparasitic and cardiac activity. The polynucleotide and
CC	XX	protein sequences are used in the diagnosis of cancer, particularly
CC	XX	breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC	XX	and antagonists may also be used in the diagnosis, prevention and treatment
CC	XX	of immune disorders e.g. Addison's disease, allergies, autoimmune
CC	XX	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC	XX	disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC	XX	cardiovascular disorders such as myocardial ischaemias; wound healing;
CC	XX	neurological diseases such as cerebral anoxia and epilepsy; and
CC	XX	infectious diseases
XX	XX	
SQ	XX	Sequence 241 AA;
XX	XX	
Query Match	100.0%; Score 82; DB 3; Length 241;	
Best Local Similarity	100.0%; Pred. No. 0.024;	
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 EEEEDDEDDEDD 15	
Db	32 EEEEEDEDEDD 46	
RESULT 6		
AAB82806		
ID AAB82806 standard; protein; 538 AA.		
XX		
AC AAB82806;		

XX 12-NOV-2001 (first entry)
 XX Human low density lipoprotein binding protein 2 (LBP-2).
 DE Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 XX atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 KW Homo sapiens.
 XX WO200164874-A2.
 XX 07-SEP-2001.
 XX 28-FEB-2001; 2001WO-US006356.
 XX 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX (BOST-) BOSTON HEART FOUND INC.
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 2001-565505/63.
 DR N-PSDB; AAH26499.
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX Claim 13(j); Fig 7A; 143pp; English.
 XX The present sequence is that of novel human low density lipoprotein
 CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the
 CC coding region of isolated genomic DNA (see AAH26499). It differs from the
 CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the
 CC presence of an additional 321 amino acids at the N-terminus (the cDNA
 CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP
 CC polypeptides of the invention that are capable of binding to native and
 CC methylated low density lipoproteins. Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs.
 CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529
 CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of
 CC determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX Sequence 538 AA;
 SQ

Query Match 100.0%; Score 82; DB 4; Length 538;
 Best Local Similarity 100.0%; Pred. No. 0.055;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEEEDDEDEDEDD 15
 |||||
 DB 329 EEEEDDEDEDEDD 343

RESULT 7
 ABB84611
 ID ABB84611 standard; protein; 358 AA.
 XX ABB84611;
 XX 24-FEB-2003 (first entry)
 XX Soybean nucleosome assembly protein 1.
 DE NARC10; NARC16; cytosolic; immunosuppressive; dermatological; cardiant;
 KW antiinflammatory; nephrotropic; anti-HIV; nontropic; neuroprotective;
 KW antianaemic; cerebroprotective; vasotropic; antidiabetic; anticonvulsant;

KW immunosuppressive; thyromimetic; immunostimulant; antianaemic; hypotensive;
 KW tranquiliser; neuroleptic; gene therapy; gene mapping; apoptosis;
 KW AIDS: cell cycle disruption; programmed cell death regulation;
 KW viral infection; nucleosome assembly; phosphate homeostasis;
 KW cell cycle regulation; cancer; follicular lymphoma; carcinoma;
 KW p53 mutation; graft rejection; hormone-dependent tumour;
 KW autoimmune disorder; valvular heart disease;
 KW systemic lupus erythematosus; diabetes; Hashimoto's thyroiditis;
 KW immune-mediated glomerulonephritis; virus-induced lymphocyte depletion;
 KW acquired immunodeficiency syndrome; neurodegenerative disease; stroke;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration;
 KW myelodysplastic syndrome; ischaemic injury; myocardial infarction;
 KW reperfusion injury; liver disease; idiopathic dilated cardiomyopathy;
 KW ischaemic cardiomyopathy; aplastic anaemia; chronic neuropania; mania;
 KW myelodysplastic syndrome; central nervous system disorder; anxiety;
 KW senile dementia; Huntington's disease; hypertension; schizophrenia;
 KW severe bipolar affective disorder; nucleosome assembly protein 1.
 OS Glycine max.
 XX WO200281516-A2.
 XX 17-OCT-2002.
 XX 16-JAN-2002; 2002WO-US001098.
 XX 16-JAN-2001; 2001US-0262306P.
 PR 15-JAN-2002; 2002US-00047855.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Chiang LW;
 XX WPI; 2003-058503/05.
 XX Novel isolated programmed cell death-related polypeptide, NARC10 and
 PT NARC16, useful for treating disorders associated with abnormal apoptotic
 PT process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.
 XX Disclosure; Fig 1A-D; 123pp; English.
 XX This invention describes novel cell death-related polypeptides NARC10 and
 CC NARC16, located on chromosome 4q11-q21 and which have cardiant, antiHIV,
 CC immunosuppressive, dermatological, antiinflammatory, cerebroprotective,
 CC nontropic, neuroprotective, antianaemic, cardiant, vasotropic, antianaemic,
 CC antidiabetic, immunosuppressive, cytosolic, thyromimetic, nephrotropic,
 CC immunostimulant, anticonvulsant, tranquiliser, hypotensive and
 CC neuroleptic activity and can be used in gene therapy. The products of the
 CC invention can be used to modulate NARC10 or NARC16 polypeptides or
 CC polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene
 CC regions associated with genetic disease or to associate NARC10 or NARC16
 CC with a disease. The polypeptides are also useful for modulating the
 CC apoptotic process, and are therefore useful for modulating, and treating
 CC disorders associated with increased apoptosis, inhibition of apoptosis or
 CC disruptions in cell cycle, for regulating cellular functions including
 CC programmed cell death, nucleosome assembly, phosphate homeostasis and the
 CC cell cycle. Preferably, the products of the invention are useful for
 CC treating disorders associated with abnormally low rate or abnormally high
 CC rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas
 CC with p53 mutations, or hormone-dependent tumours, autoimmune disorders
 CC including systemic lupus erythematosus, diabetes, graft rejection,
 CC Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral
 CC infections e.g. infections caused by herpes viruses, virus-induced
 CC lymphocyte depletion (including acquired immunodeficiency syndrome
 CC (AIDS)), neurodegenerative diseases manifested by loss of specific sets
 CC of neurons (including Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis
 CC pigmentosa, and cerebellar degeneration), myelodysplastic syndromes
 CC (including aplastic anaemia), ischaemic injuries (including myocardial
 CC infarction, stroke and reperfusion injury), and toxin (e.g. alcohol)
 CC induced liver disease, idiopathic dilated cardiomyopathy, ischaemic
 CC cardiomyopathy and valvular heart disease, aplastic anaemia, chronic

CC neutropaenia, and myelodysplastic syndromes, central nervous system
 CC disorders, senile dementia, Huntington's disease, hypertension,
 CC schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar
 CC affective disorder (BP-I). This sequence represents the soybean
 CC nucleosome assembly protein 1 described in the method of the invention
 XX
 XX Sequence 358 AA;

Query Match 91.5%; Score 75; DB 6; Length 358;
 Best Local Similarity 86.7%; Pred. No. 0.24;
 Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDD 15
 :||||:|||||
 Db 307 DEEDEDEDEDD 321

RESULT 8
 AAY50875
 ID AAY50875 standard; protein; 457 AA.

XX AC AAY50875;
 XX 24-FEB-2000 (first entry)
 XX Murine NNX3 protein.
 DE NNX3; murine; cytostatic; marker; lung tumor; Hodgkin's disease;
 XX treatment.
 KW Mus musculus.
 OS WO9960115-A2.
 PN 25-NOV-1999.
 PD 17-MAY-1999; 99WO-EP003374.
 XX 18-MAY-1998; 98EP-00201642.
 PR (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
 PA Van Leuven F;

PI WPI; 2000-053295/04.
 DR N-PSDB; AAZ43754.
 XX Novel proteins useful as markers and for treating lung tumors and/or
 PT Hodgkin's disease.
 PS Claim 2; Page 47-49; 50pp; English.

XX This invention describes two novel NNX3 proteins isolated from human and
 CC murine sources which have cytostatic activity. The NNX3 mRNA and/or
 CC protein is useful as a marker for lung tumors or Hodgkin's disease. The
 CC proteins form pharmaceutical compositions useful for treating lung tumors
 CC and/or Hodgkin's disease. Polynucleotide products of the invention form
 CC an assay for screening the expression of these nucleic acids. Antibodies
 CC raised against the proteins of the invention form an assay for detecting
 CC the proteins. The probes form an assay for detecting and/or amplifying
 CC NNX3 polynucleotides. The expression of NNX3 in humans and mice suggest
 CC that NNX3 is useful as a marker for lung tumors and Hodgkin's disease.
 CC Therefore, the protein facilitates therapies for these conditions. This
 CC sequence represents the murine NNX3 protein described in the method of
 CC the invention

XX Sequence 457 AA;

Query Match 89.0%; Score 73; DB 3; Length 457;
 Best Local Similarity 80.0%; Pred. No. 0.52;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDD 15

Db 229 EEEDEDEDEDD 243
 :||||:||||:|
 RESULT 9
 AAR95607
 ID AAR95607 standard; protein; 1085 AA.

XX AC AAR95607;
 XX 12-OCT-1996 (first entry)
 DT RRP3 telomerase-associated protein.
 DE Yeast; RRP3 protein; telomerase-associated protein; STR7;
 XX suppressor of telomeric repression-7; telomerase; ribonucleoprotein;
 KW telomere; tumour; pathogen; sperm; ovum; reporter gene; drug screening;
 KW antibody; immunoassay; antitumour; antiseptic; contraceptive;
 KW infertility; diagnostic; gene therapy.
 XX Saccharomyces cerevisiae.

XX WO9612811-A2.

XX 02-MAY-1996.
 XX 20-OCT-1995; 95WO-US013801.
 XX 20-OCT-1994; 94US-00326781.
 PR 28-APR-1995; 95US-00431080.

XX (ARCH-) ARCH DEV CORP.

XX Gottschling DE, Singer MS;

XX WPI; 1996-239169/24.

DR N-PSDB; AAT27052.

XX Novel telomerase associated polypeptide(s) and related nucleic acid -
 PT useful for detecting e.g. tumour cells or pathogens.
 XX Example 10; Page 294-302; 349pp; English.

PS RRP3 telomerase-associated protein (containing fragment STR7 (AAR95606)
 XX is found in conjunction with Saccharomyces cerevisiae telomerase, a
 CC ribonucleoprotein required for telomere replication. Other proteins
 CC associated with telomerase are given in AAR95601-05 (Other STR proteins).
 CC These proteins combine with telomerase to repress telomere silencing of
 CC gene expression. Oligonucleotides from the encoding sequence may be used
 CC to detect non-ciliate telomerase-associated genes, e.g. in tumour,
 CC pathogen, sperm or ovum cells. New telomerase-associated sequences may be
 CC detected by a reporter gene expression system linked to an expression-
 CC repressing telomerase sequence, and binding compounds, e.g. antibodies, may
 CC be detected by complex formation with telomerase components. The products
 CC may be used as antitumour, antiseptic or contraceptive agents, in
 CC infertility diagnosis, or in gene therapy

XX Sequence 1085 AA;

Query Match 89.0%; Score 73; DB 2; Length 1085;
 Best Local Similarity 80.0%; Pred. No. 1.2;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDD 15

Db 133 EEEDEDEDEDD 147

RESULT 10
 ADK62464
 ID ADK62464 standard; protein; 1085 AA.

XX AC ADK62464;

XX 06-MAY-2004 (first entry)
 DT Disease treating protein complex-derived protein #358.
 XX protein complex; drug target; diagnosis.

XX Unidentified.

OS EP1338608-A2.
 XX 27-AUG-2003.

XX 20-DEC-2002; 2002EP-00102902.
 XX 20-DEC-2001; 2001EP-00130253.

XX (CELL-) CELLZONE AG.
 XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;
 PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;
 PI Michon A, Leutwein C, Rick J;

XX WPI; 2003-638460/61.
 DR N-PSDB; ADK62465.

XX New proteins and protein complexes from eukaryotes, useful as targets in
 PT drug screening, or in diagnosing or screening for the presence of a
 PT disease or disorder, or a predisposition for developing a disease or
 PT disorder in a subject.

XX Disclosure; SEQ ID NO 715; 13pp; English.

XX The invention relates to novel protein complexes comprising a first and a
 CC second protein, or its derivative, fragment, homologue or variant. The
 CC proteins are selected from given protein complexes, which are not defined
 CC in the specification. The variants are encoded by nucleic acids that
 CC hybridize to the nucleic acids encoding the proteins under low stringency
 CC conditions. The protein complexes are useful as targets for an active
 CC agent of a pharmaceutical. These protein complexes are particularly
 CC useful as drugs targets for the treatment or preventing of a disease or
 CC disorder. The complexes and methods above are useful in diagnosing or
 CC screening for the presence of a disease or disorder or a predisposition
 CC for developing a disease or disorder in a subject. These are also useful
 CC in screening for a drug for treatment or prevention of a disease or
 CC disorder. The molecule that modulates the amount, activity or protein
 CC components of the complex is useful for the manufacture of a medicament
 CC for the treatment or prevention of a disease or disorder. This sequence
 CC corresponds to a protein of the invention. (Note: the sequence data for
 CC this patent did not form part of the printed specification but was
 CC obtained from the EPO in electronic format).

XX SQ Sequence 1085 AA;

Query Match 89.0%; Score 73; DB 7; Length 1085;
 Best Local Similarity 80.0%; Pred. No. 1-2;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||:|:|:|:
 Db 133 EEEEDDEDEDDDD 147

RESULT 11
 ABG93347
 ID ABG93347 standard; protein; 114 AA.

XX AC ABG93347;
 XX 21-NOV-2002 (first entry)

XX C. albicans BAX-associated protein fragment SEQ ID 652.
 DE
 XX

KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death.

XX Candida albicans.
 XX WO200264766-A2.

XX 22-AUG-2002.

XX 21-DEC-2001; 2001WO-EP015398.

XX 22-DEC-2000; 2000EP-00870318.

XX 04-JAN-2001; 2001EP-00870002.

XX 09-JAN-2001; 2001EP-00870003.

XX (JANC) JANSSEN PHARM NV.

XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
 PI WPI; 2002-667002/71.

XX DR N-PSDB; ABQ76613.

XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.

XX Claim 36; Fig 2; 344pp; English.

XX This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cytostatic, fungicide; immunosuppressive, virucide and
 CC vasotropic activity and can be used in vaccines or for gene therapy. The
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
 CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenous flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polypeptide associated with the Bax gene
 CC described in the disclosure of the invention

XX SQ Sequence 114 AA;

Query Match 87.8%; Score 72; DB 5; Length 114;
 Best Local Similarity 80.0%; Pred. No. 0.17;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||:|:|:|:
 Db 22 EEEEDDEDEDEDD 36

RESULT 12
 AAG29185
 ID AAG29185 standard; protein; 178 AA.

XX AC AAG29185;
 XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 34681.
 DE
 XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01231180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 08-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.

PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 08-AUG-1999; 99US-0147192P.
PR 08-AUG-1999; 99US-0147262P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 15-AUG-1999; 99US-0148684P.
PR 15-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 87.8%; Score 72; DB 3; Length 178;
Best Local Similarity 80.0%; Pred. No. 0.27;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDD 15
|||||:|:|:|
Db 164 EEEEDDDDDDEED 178

RESULT 14
ADN74341
ID ADN74341 standard; protein; 178 AA.
XX
AC ADN74341;
XX

DT 15-JUL-2004 (first entry)
XX

DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 2236.
XX
KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
KW animal feed product; thale cress; cell wall biosynthesis;
KW nitrogen metabolism; carbon metabolism.

OS Arabidopsis thaliana.
XX

PN WO2004035798-A2.
XX

XX 29-APR-2004.
PD

XX 20-OCT-2003; 2003WO-EP011658.
PF

XX 18-OCT-2002; 2002EP-00079408.
PR

XX

PA (CROP-) CROPDESIGN NV.
 XX Inze D, De Veylder L, Vlieghe K;
 XX WPI: 2004-348466/32.
 DR N-PSDB; ADN74340.
 XX
 XX Altering plant characteristics, useful for producing plants for enzyme or
 PT pharmaceutical production comprises modifying in a plant, expression of
 PT one or more nucleic acids and/or modifying level or activity of one or
 PT more proteins.
 XX
 XX Claim 1; SEQ ID NO 2236; 134pp; English.
 PS
 XX This invention relates to a novel method for altering one or more plant
 CC characteristics. Specifically, it refers to identifying genes that are up
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
 CC alter plant characteristics accordingly. The present invention describes
 CC generating transgenic plants for the production of growth regulators,
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
 CC the altered plant characteristics are selected from increased yield or
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture
 CC or physiology, altered endoreduplication, biochemistry, signal
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,
 CC each relative to the corresponding wild type plants. Accordingly, these
 CC sequences can also be useful as positive or negative selectable markers
 CC during transformation of cells or tissues. The identified genes play a
 CC role in a variety of biological processes such as DNA replication, cell
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polypeptide sequence is thale cress protein
 CC expressed by a gene repressed 1.3 fold or more in plants overexpressing
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the
 CC invention.
 XX
 XX SQ Sequence 178 AA;
 Query Match 87.8%; Score 72; DB 8; Length 178;
 Best Local Similarity 80.0%; Pred. No. 0.27;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEEEDDEDEDEDD 15
 Db 164 EEEEDDDDDDEED 178
 |||||:|:|:|:
 RESULT 15
 ADN73879
 ID ADN73879 standard; protein; 178 AA.
 XX
 XX AC ADN73879;
 XX
 XX DT 15-JUL-2004 (first entry)
 XX
 XX DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 1774.
 XX
 XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
 KW animal feed product; thale cress; cell wall biosynthesis;
 KW nitrogen metabolism; carbon metabolism.
 XX
 XX OS Arabidopsis thaliana.
 XX
 XX PN WO2004035798-A2.
 XX
 XX PD 29-APR-2004.
 XX
 XX PF 20-OCT-2003; 2003WO-EP011658.
 XX
 XX PR 18-OCT-2002; 2002EP-00079408.
 XX
 XX PA (CROP-) CROPDESIGN NV.
 XX Inze D, De Veylder L, Vlieghe K;
 PI

XX WPI: 2004-348466/32.
 DR N-PSDB; ADN73878.
 XX
 XX Altering plant characteristics, useful for producing plants for enzyme or
 PT pharmaceutical production comprises modifying in a plant, expression of
 PT one or more nucleic acids and/or modifying level or activity of one or
 PT more proteins.
 XX
 XX Claim 1; SEQ ID NO 1774; 134pp; English.
 PS
 XX This invention relates to a novel method for altering one or more plant
 CC characteristics. Specifically, it refers to identifying genes that are up
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
 CC alter plant characteristics accordingly. The present invention describes
 CC generating transgenic plants for the production of growth regulators,
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
 CC the altered plant characteristics are selected from increased yield or
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture
 CC or physiology, altered endoreduplication, biochemistry, signal
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,
 CC each relative to the corresponding wild type plants. Accordingly, these
 CC sequences can also be useful as positive or negative selectable markers
 CC during transformation of cells or tissues. The identified genes play a
 CC role in a variety of biological processes such as DNA replication, cell
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polypeptide sequence is thale cress protein
 CC expressed by a gene repressed 1.3 fold or more in plants overexpressing
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the
 CC invention.
 XX
 XX SQ Sequence 178 AA;
 Query Match 87.8%; Score 72; DB 8; Length 178;
 Best Local Similarity 80.0%; Pred. No. 0.27;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEEEDDEDEDEDD 15
 Db 164 EEEEDDDDDDEED 178
 |||||:|:|:|:
 Search completed: September 20, 2005, 12:34:42
 Job time : 10.9229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 2.27662 Seconds
(without alignments)
491.841 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82
Sequence: 1 EEEEDDEDEEDD 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	15	3	US-08-979-608A-19
2	82	100.0	15	4	US-09-517-849-19
3	82	100.0	15	4	US-09-616-289-19
4	82	100.0	26	3	US-08-979-608A-20
5	82	100.0	26	4	US-09-517-849-20
6	82	100.0	26	4	US-09-616-289-20
7	82	100.0	217	3	US-08-979-608A-7
8	82	100.0	217	4	US-09-517-849-7
9	82	100.0	217	4	US-09-616-289-7
10	82	100.0	538	4	US-09-616-289-43
11	73	89.0	1085	1	US-08-431-080-28
12	73	89.0	1085	2	US-08-938-534-28
13	73	89.0	1085	3	US-09-345-294-28
14	72	87.8	116	4	US-09-248-796A-17428
15	72	87.8	428	4	US-09-248-796A-15197
16	71	86.6	213	4	US-09-949-016-10813
17	71	86.6	667	4	US-09-538-092-1307
18	71	86.6	1972	4	US-09-418-710-21
19	71	86.6	1972	4	US-09-839-479-21
20	70	85.4	209	4	US-09-214-881A-8
21	70	85.4	214	4	US-09-538-092-883
22	70	85.4	214	4	US-09-214-881A-1
23	70	85.4	226	1	US-08-431-080-26
24	70	85.4	226	2	US-08-938-534-26
25	70	85.4	226	3	US-09-345-294-26
26	70	85.4	706	4	US-09-538-092-957
27	70	85.4	747	4	US-09-949-016-10040

28	70	85.4	932	4	US-09-248-796A-19128	Sequence 19128, A
29	69	84.1	279	4	US-09-699-266A-7	Sequence 7, Appli
30	69	84.1	352	1	US-08-552-142A-4	Sequence 4, Appli
31	69	84.1	352	1	US-08-910-973-4	Sequence 4, Appli
32	69	84.1	352	4	US-09-499-227-4	Sequence 4, Appli
33	69	84.1	352	5	PCT-US95-05741-4	Sequence 4, Appli
34	69	84.1	700	2	US-08-568-459A-10	Sequence 10, Appl
35	69	84.1	700	2	US-08-487-826B-10	Sequence 10, Appl
36	69	84.1	700	3	US-09-210-288-10	Sequence 10, Appl
37	69	84.1	2182	2	US-08-487-826B-16	Sequence 16, Appl
38	68	82.9	764	4	US-09-370-838-67	Sequence 67, Appl
39	68	82.9	764	4	US-09-538-092-944	Sequence 944, App
40	68	82.9	764	4	US-09-854-133-67	Sequence 67, Appl
41	68	82.9	781	4	US-09-949-016-9773	Sequence 9773, Ap
42	68	82.9	1162	2	US-08-728-323A-2	Sequence 2, Appli
43	68	82.9	1162	3	US-09-298-568-2	Sequence 2, Appli
44	68	82.9	1162	4	US-09-410-399-2	Sequence 2, Appli
45	68	82.9	1162	4	US-09-894-273-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-979-608A-19
; Sequence 19, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-979-608A-19

Query Match 100.0%; Score 82; DB 3; Length 15;

```
Best Local Similarity 100.0%; Pred. No. 0.00065; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

QY 1 EEEEDDEDEDEDD 15
   |||||
Db 1 EEEEDDEDEDEDD 15

RESULT 2
US-09-517-849-19
; Sequence 19, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-517-849-19

Query Match 100.0%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
   |||||
Db 1 EEEEDDEDEDEDD 15

RESULT 3
US-09-616-289-19
; Sequence 19, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-19

Query Match 100.0%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
   |||||
Db 1 EEEEDDEDEDEDD 15

RESULT 4
US-08-979-608A-20
; Sequence 20, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
```

```
/
/ LENGTH: 26 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-979-608A-20
Query Match 100.0%; Score 82; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. NO. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 5
US-09-517-849-20
; Sequence 20, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION NUMBER: US/09/517,849
; APPLICATION DATE: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20
Query Match 100.0%; Score 82; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. NO. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 6
US-09-616-289-20
; Sequence 20, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; BINDING PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-616-289-20
Query Match 100.0%; Score 82; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. NO. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 7
US-08-979-608A-7
; Sequence 7, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
```

; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match 100.0%; Score 82; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
DB 8 EEEEDDEDEDD 22

RESULT 8
US-09-517-849-7
; Sequence 7, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-7

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match 100.0%; Score 82; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.0083;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
DB 8 EEEEDDEDEDD 22

RESULT 9
US-09-616-289-7
; Sequence 7, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-7

Query Match 100.0%; Score 82; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
DB 8 EEEEDDEDEDD 22

RESULT 10
US-09-616-289-43
; Sequence 43, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43

```
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-43

Query Match      100.0%; Score 82; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDD 15
    |||||
Db 329 EEEEDDEDEDD 343

RESULT 11
US-08-431-080-28
; Sequence 28, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,080
; FILING DATE:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; APPLICATION DATA:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-431-080-28

Query Match      89.0%; Score 73; DB 1; Length 1085;
Best Local Similarity 80.0%; Pred. No. 0.39;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDD 15
    |||||
Db 133 EEEEDDEDDDD 147

RESULT 13
US-09-345-294-28
; Sequence 28, Application US/09345294
; Patent No. 6387619
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/345,294
```

```
; FILING DATE: 30-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,080
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-345-294-28

Query Match      89.0%; Score 73; DB 3; Length 1085;
Best Local Similarity 80.0%; Pred. No. 0.39;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDEDD 15
Db      133 EEEEDDEDEDEDD 147

RESULT 14
US-09-248-796A-17428
; Sequence 17428, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17428
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17428

Query Match      87.8%; Score 72; DB 4; Length 116;
Best Local Similarity 80.0%; Pred. No. 0.061;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDEDD 15
Db      24 EEEEDDEDEDEDD 38

RESULT 15
US-09-248-796A-15197
; Sequence 15197, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
```

```
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15197
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15197

Query Match      87.8%; Score 72; DB 4; Length 428;
Best Local Similarity 80.0%; Pred. No. 0.21;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDEDD 15
Db      135 EEEEDDEDEDEDD 149

Search completed: September 20, 2005, 12:45:29
Job time : 3.27662 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 3.94614 Seconds
(without alignments)
491.841 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 ESEBDDDEDEDDVSGSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	26	3	US-08-979-608A-20
2	136	100.0	26	4	US-09-517-849-20
3	136	100.0	26	4	US-09-616-289-20
4	136	100.0	217	3	US-08-979-608A-7
5	136	100.0	217	4	US-09-517-849-7
6	136	100.0	217	4	US-09-616-289-7
7	136	100.0	538	4	US-09-616-289-43
8	104.5	76.8	28	3	US-08-979-608A-26
9	104.5	76.8	28	4	US-09-517-849-26
10	104.5	76.8	28	4	US-09-616-289-26
11	104.5	76.8	232	3	US-08-979-608A-3
12	104.5	76.8	232	4	US-09-517-849-3
13	104.5	76.8	232	4	US-09-616-289-3
14	104.5	76.8	252	3	US-08-979-608A-4
15	104.5	76.8	252	4	US-09-517-849-4
16	104.5	76.8	252	4	US-09-616-289-4
17	104.5	76.8	317	3	US-08-979-608A-2
18	104.5	76.8	317	4	US-09-517-849-2
19	104.5	76.8	317	4	US-09-616-289-2
20	104.5	76.8	550	4	US-09-616-289-47
21	88	64.7	162	1	US-07-879-685B-4
22	88	64.7	599	4	US-09-538-092-864
23	83.5	61.4	305	4	US-09-645-337A-8
24	83	61.0	428	4	US-09-248-796A-15197
25	82	60.3	15	3	US-08-979-608A-19
26	82	60.3	15	4	US-09-517-849-19
27	82	60.3	15	4	US-09-616-289-19

28	81	59.6	557	4	US-09-248-796A-19073	Sequence 19073, A
29	81	59.6	1162	2	US-08-728-323A-2	Sequence 2, Appli
30	81	59.6	1162	3	US-09-298-568-2	Sequence 2, Appli
31	81	59.6	1162	4	US-09-410-399-2	Sequence 2, Appli
32	81	59.6	1162	4	US-09-894-273-2	Sequence 2, Appli
33	80	58.6	706	4	US-09-538-092-957	Sequence 957, App
34	80	58.8	747	4	US-09-949-016-10040	Sequence 10040, A
35	79	58.1	764	1	US-08-375-300-4	Sequence 4, Appli
36	79	58.1	764	3	US-09-177-431-4	Sequence 4, Appli
37	79	58.1	764	5	PCT-US95-16930-4	Sequence 2, Appli
38	79	58.1	1089	1	US-08-375-300-2	Sequence 2, Appli
39	79	58.1	1089	3	US-09-177-431-2	Sequence 2, Appli
40	79	58.1	1089	5	PCT-US95-16930-2	Sequence 2, Appli
41	79	58.1	1972	4	US-09-418-710-21	Sequence 21, Appl
42	79	58.1	1972	4	US-09-839-479-21	Sequence 21, Appl
43	78	57.4	932	4	US-09-248-796A-19128	Sequence 19128, A
44	77	56.6	116	4	US-09-248-796A-17428	Sequence 17428, A
45	76.5	56.2	1085	1	US-08-431-080-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-08-979-608A-20
; Sequence 20, Application US/08979608A
; Patent No. 6355451

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M. S.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq For Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/979,608A

; FILING DATE: 26-NOV-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,547

; FILING DATE: 03-JUN-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 26 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-08-979-608A-20

Query Match 100.0%; Score 136; DB 3; Length 26;

```

Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGSEVPESD 26
   |||||||
Db 1 EEEEDDEDEDEDDVSEGSEVPESD 26

RESULT 2
US-09-517-849-20
; Sequence 20, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
; US-09-517-849-20

Query Match 100.0%; Score 136; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGSEVPESD 26
   |||||||
Db 1 EEEEDDEDEDEDDVSEGSEVPESD 26

RESULT 3
US-09-616-289-20
; Sequence 20, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

```


;
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match 100.0%; Score 136; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDDVSGSEVPESD 26
Db 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 5

US-09-517-849-7
; Sequence 7, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match 100.0%; Score 136; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDDVSGSEVPESD 26
Db 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 6

US-09-616-289-7
; Sequence 7, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-7

Query Match 100.0%; Score 136; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDDVSGSEVPESD 26
Db 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 7

US-09-616-289-43
; Sequence 43, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-43

Query Match 100.0%; Score 136; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 EEEEDDEDEDDVSEGSVEPESD 26
Db      329 EEEEDDEDEDDVSEGSVEPESD 354

RESULT 8
US-08-979-608A-26
; Sequence 26, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-979-608A-26

Query Match      76.8%; Score 104.5; DB 3; Length 28;
Best Local Similarity 74.1%; Pred. No. 1.2e-06;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY      1 EEEEDDEDEDDVSEGSVEPESD 26
Db      2 EEEEDDEDEDDVSEGSVEPESD 28

RESULT 9
US-09-517-849-26
; Sequence 26, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26

QY      1 EEEEDDEDEDDVSEGSVEPESD 26
Db      2 EEEEDDEDEDDVSEGSVEPESD 28

RESULT 10
US-09-616-289-26
; Sequence 26, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26

QY      1 EEEEDDEDEDDVSEGSVEPESD 26
Db      2 EEEEDDEDEDDVSEGSVEPESD 28

Query Match      76.8%; Score 104.5; DB 4; Length 28;
Best Local Similarity 74.1%; Pred. No. 1.2e-06;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;
```

```
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-26

Query Match          76.8%; Score 104.5; DB 4; Length 28;
Best Local Similarity 74.1%; Pred. No. 1.1e-06;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDDEDEDD-VSEGSEVPESD 26
   |||:|||||:|||||
Db 2 EEEEEDEDDDDVVSEGSEVPESD 28
   |||:|||||:|||||

RESULT 11
US-08-979-608A-3
; Sequence 3, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match          76.8%; Score 104.5; DB 3; Length 232;
Best Local Similarity 74.1%; Pred. No. 1.1e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDDEDEDD-VSEGSEVPESD 26
   |||:|||||:|||||
Db 2 EEEEEDEDDDDVVSEGSEVPESD 47
   |||:|||||:|||||

RESULT 12
US-08-979-608A-3
; Sequence 3, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match          76.8%; Score 104.5; DB 3; Length 232;
Best Local Similarity 74.1%; Pred. No. 1.1e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDDEDEDD-VSEGSEVPESD 26
   |||:|||||:|||||
Db 2 EEEEEDEDDDDVVSEGSEVPESD 47
   |||:|||||:|||||

RESULT 13
US-09-616-289-3
; Sequence 3, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
```

```
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-3

Query Match          76.8%; Score 104.5; DB 4; Length 232;
Best Local Similarity 74.1%; Pred. No. 1.1e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDD-VSEGSVPESD 26
   |||:::|||||
Db 21 EEEEDDEDDDDVVSEGSVPESD 47
   |||:::|||||

RESULT 14
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION NUMBER: US/09/517,849
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match          76.8%; Score 104.5; DB 3; Length 252;
Best Local Similarity 74.1%; Pred. No. 1.2e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDD-VSEGSVPESD 26
   |||:::|||||
Db 21 EEEEDDEDDDDVVSEGSVPESD 47
   |||:::|||||

RESULT 15
US-09-517-849-4
; Sequence 4, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION NUMBER: US/09/517,849
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-517-849-4

Query Match          76.8%; Score 104.5; DB 4; Length 252;
Best Local Similarity 74.1%; Pred. No. 1.2e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDD-VSEGSVPESD 26
   |||:::|||||
Db 41 EEEEDDEDDDDVVSEGSVPESD 67
   |||:::|||||

Search completed: September 20, 2005, 12:45:29
Job time : 3.94614 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 8.28029 Seconds
(without alignments)
733.538 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82

Sequence: 1 EEEEDDEDEDD 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 segs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	9	US-09-962-055-19
2	82	100.0	15	9	US-09-976-740-19
3	82	100.0	15	13	US-10-023-529-19
4	82	100.0	15	13	US-10-023-523-19
5	82	100.0	15	15	US-10-616-187-19
6	82	100.0	15	15	US-10-671-242-19
7	82	100.0	26	9	US-09-962-055-20
8	82	100.0	26	9	US-09-976-740-20
9	82	100.0	26	13	US-10-023-529-20
10	82	100.0	26	13	US-10-023-523-20
11	82	100.0	26	15	US-10-616-187-20
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 20, Appl
					Sequence 20, Appl
					Sequence 20, Appl
					Sequence 20, Appl
					Sequence 20, Appl

12	100.0	26	15	US-10-671-242-20	Sequence 20, Appl
13	100.0	217	9	US-09-962-055-7	Sequence 7, Appl
14	100.0	217	9	US-09-976-740-7	Sequence 7, Appl
15	100.0	217	13	US-10-023-529-7	Sequence 7, Appl
16	100.0	217	13	US-10-023-523-7	Sequence 7, Appl
17	100.0	217	15	US-10-616-187-7	Sequence 7, Appl
18	100.0	217	15	US-10-671-242-7	Sequence 7, Appl
19	100.0	241	9	US-09-925-298-665	Sequence 665, App
20	100.0	241	14	US-10-103-806-665	Sequence 665, App
21	100.0	538	9	US-09-976-740-43	Sequence 43, Appl
22	100.0	538	13	US-10-023-529-43	Sequence 43, Appl
23	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
24	100.0	538	15	US-10-616-187-43	Sequence 43, Appl
25	100.0	538	15	US-10-671-242-43	Sequence 43, Appl
26	91.5	358	14	US-10-047-855-9	Sequence 9, Appl
27	91.5	358	15	US-10-424-599-172634	Sequence 172634, A
28	91.5	395	15	US-10-425-114-40855	Sequence 40855, A
29	89.0	123	16	US-10-425-115-307016	Sequence 307016, A
30	87.8	80	16	US-10-767-701-48120	Sequence 48120, A
31	87.8	114	16	US-10-451-467A-652	Sequence 652, App
32	87.8	181	15	US-10-425-114-72908	Sequence 72908, A
33	87.8	288	15	US-10-369-493-6891	Sequence 6891, Ap
34	87.8	427	14	US-10-032-585-7386	Sequence 7386, Ap
35	86.6	166	15	US-10-424-599-169943	Sequence 169943, A
36	86.6	167	16	US-10-739-930-9078	Sequence 9078, Ap
37	86.6	376	15	US-10-072-012-653	Sequence 653, App
38	86.6	474	15	US-10-072-012-651	Sequence 651, App
39	86.6	597	15	US-10-331-496A-39	Sequence 39, Appl
40	86.6	612	15	US-10-072-012-650	Sequence 650, App
41	86.6	667	15	US-10-072-012-649	Sequence 649, App
42	86.6	667	16	US-10-723-860-823	Sequence 823, App
43	86.6	668	15	US-10-072-012-242	Sequence 242, App
44	86.6	1972	9	US-09-839-479-21	Sequence 21, Appl
45	86.6	1972	15	US-10-376-537-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-09-962-055-19

; Sequence 19, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; City: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-962-055-19

Query Match      100.0%; Score 82; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 2
US-09-976-740-19
; Sequence 19, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-19

Query Match      100.0%; Score 82; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 3
US-10-023-529-19
; Sequence 19, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-19

Query Match      100.0%; Score 82; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-19

Query Match      100.0%; Score 82; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 4
US-10-023-523-19
; Sequence 19, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-19

Query Match      100.0%; Score 82; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15
```

```
RESULT 5
US-10-616-187-19
; Sequence 19, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-19

Query Match      100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 6
US-10-671-242-19
; Sequence 19, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-19

Query Match      100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 7
US-09-962-055-20
; Sequence 20, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 82; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 8
US-09-976-740-20
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-19
```

```
Query Match      100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15
```

```
RESULT 7
US-09-962-055-20
; Sequence 20, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20
```

```
Query Match      100.0%; Score 82; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15
```

```
RESULT 8
US-09-976-740-20
```

```
; Sequence 20, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-20

Query Match      100.0%; Score 82; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 9
US-10-023-529-20
; Sequence 20, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 82; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 11
US-10-616-187-20
; Sequence 20, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 82; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 10
US-10-023-523-20
; Sequence 20, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 82; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 11
US-10-616-187-20
; Sequence 20, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 82; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15
```



```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-20

Query Match      100.0%; Score 82; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 12
US-10-671-242-20
; Sequence 20, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671.242
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-03-02
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-20

Query Match      100.0%; Score 82; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 13
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US2002005203A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962.055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match      100.0%; Score 82; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 8 EEEEDDEDEDEDD 22

RESULT 14
US-09-976-740-7
; Sequence 7, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-976-740-7

Query Match      100.0%; Score 82; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEEDEDEDD 15
Db 8 EEEEEDEDEDD 22

RESULT 15
US-10-023-529-7
; Sequence 7, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 82; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEEDEDEDD 15
Db 8 EEEEEDEDEDD 22

Search completed: September 20, 2005, 12:53:06
Job time : 9.28029 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 7.61934 Seconds
(without alignments)
1008.117 Million cell updates/sec

Title: US-10-671-242-19
Perfect score: 82
Sequence: 1 EEEEDDEDEEDD 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	236	2 Q61M4	Q61m4 homo sapien
2	82	100.0	285	2 Q6P0R3	Q6p0r3 homo sapien
3	82	100.0	295	2 Q6P1S7	Q6p1s7 homo sapien
4	82	100.0	538	2 Q6SPF0	Q6spf0 homo sapien
5	79	96.3	239	1 CENB_SHEEP	F49451 ovis aries
6	76	92.7	2246	2 Q8T253	Q8t253 dictyosteli
7	75	91.5	358	2 Q3R892	Q3r892 glycine max
8	75	91.5	1260	2 Q7RKZ2	Q7rkz2 plasmodium
9	74	90.2	694	1 NUCL_CHICK	P15771 gallus gall
10	73	89.0	370	2 Q8BWG3	Q8bwg3 mus musculu
11	73	89.0	409	2 Q8C182	Q8c182 mus musculu
12	73	89.0	457	2 Q9Z243	Q9z243 mus musculu
13	73	89.0	625	2 Q6GNP2	Q6gnp2 xenopus lae
14	73	89.0	1085	1 IFH1_YEAST	P39520 saccharomyc
15	73	89.0	1359	2 Q7S403	Q7s403 neurospora
16	73	89.0	2404	2 Q6L876	Q6l876 plasmodium
17	73	89.0	3933	2 Q97239	Q97239 plasmodium
18	72	87.8	178	2 Q49595	Q49595 arabidopsis
19	72	87.8	215	2 Q8ITG9	Q8itg9 biophalari
20	72	87.8	287	2 Q8I4H9	Q8i4h9 caenorhabdi
21	72	87.8	288	1 SDC_CAREL	P50605 caenorhabdi
22	72	87.8	374	2 Q70Z19	Q70z19 nicotiana t
23	72	87.8	734	2 Q8IDC4	Q8idc4 plasmodium
24	72	87.8	785	2 Q9C702	Q9c702 arabidopsis
25	72	87.8	831	2 Q7RIE7	Q7rie7 plasmodium
26	72	87.8	1325	2 Q8IAR4	Q8iar4 plasmodium
27	72	87.8	1978	2 Q7RHX6	Q7rhx6 plasmodium
28	72	87.8	2104	2 Q8IKX8	Q8ikx8 plasmodium
29	71	86.6	122	2 Q6CQM0	Q6cqmq kluyveromyc
30	71	86.6	123	2 Q6FVV9	Q6fvv9 candida gla
31	71	86.6	126	2 Q75AP4	Q75ap4 ashbya goss

32	71	86.6	137	2 Q6COH4	Q6coh4 yarrowia li
33	71	86.6	198	1 VH23_SFVKA	Q9q907 shope fibro
34	71	86.6	211	1 HMI_X_HUMAN	Q9ugv6 homo sapien
35	71	86.6	222	2 Q6PU84	Q6pu84 branchiosto
36	71	86.6	270	2 Q80VC5	Q80vc5 mus musculu
37	71	86.6	373	2 Q9BR60	Q9br60 homo sapien
38	71	86.6	380	2 Q7TNV0	Q7tnv0 mus musculu
39	71	86.6	380	2 Q8BZV6	Q8bzv6 mus musculu
40	71	86.6	393	2 Q7S6W0	Q7s6w0 neurospora
41	71	86.6	612	2 Q61BL2	Q6ibl2 homo sapien
42	71	86.6	612	2 Q9BW99	Q9bw99 homo sapien
43	71	86.6	667	1 OS9_HUMAN	Q13438 homo sapien
44	71	86.6	712	1 NUCL_RAT	P13383 rattus norv
45	71	86.6	727	2 Q60460	Q60460 cricetus

ALIGNMENTS

RESULT 1					
Q96IM4					
ID	Q96IM4	PRELIMINARY;	PRT;	236 AA.	
AC	Q96IM4;				
DT	01-DEC-2001 (TREMBlrel. 19, Created)				
DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)				
DT	25-OCT-2004 (TREMBlrel. 28, Last annotation update)				
DE	LOC90378 protein (Hypothetical protein) (Fragment).				
GN	Names=LOC90378;				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA	Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RA	Strausberg R.;				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RA	Director MGC Project;				
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC007384; AAH07384.2; -				
DR	EMBL; BC080588; AAH0588.1; -				
DR	HSSP; P39769; 1KW4				
DR	InterPro; IPR001660; SAM				
DR	InterPro; IPR010993; SAM_homology				
DR	Pfam; PF00536; SAM_1; 1.				
DR	SMART; SM00454; SAM; 1.				

```

DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 82; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 27 EEEEDDEDEDEDD 41

RESULT 2
Q6P0R3 PRELIMINARY; PRT; 285 AA.
ID Q6P0R3
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF1F901 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 76 EEEEDDEDEDEDD 90

```

```

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
ID Q6PIS7
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC90378 protein (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 86 EEEEDDEDEDEDD 100

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
ID Q6SPF0
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 82; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
Db 329 EEEEDDEDEDD 343

RESULT 5
CENB_SHEEP STANDARD; PRT; 239 AA.
AC P49451;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Major centromere autoantigen B (Centromere protein B) (CENP-B) (Fragment).
DE (Fragment).
GN Name=CENPB;
OS Ovis aries (sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Burkin D.J., Jones C.A., Burkin H.R., McGrew J.A., Broad T.E.;
RX MEDLINE=97049078; PubMed=8893808;
RT "Sheep CENPB and CNPC genes show a high level of sequence similarity and conserved synteny with their human homologs.";
RL Cytogenet. Cell Genet. 74:86-89(1996).
CC -!- FUNCTION: Interacts with centromeric heterochromatin in chromosomes and binds to a specific subset of alphoid satellite DNA, called the CENP-B box. May organize arrays of centromere satellite DNA into a higher order structure which then directs centromere formation and kinetochore assembly in mammalian chromosomes (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 CENPB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U35655; AAA79098.1; -.
DR InterPro; IPR004875; CENP-B.
KW Centromere; Chromosomal protein; DNA-binding; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 239 AA; 26436 MW; 259C6C72E7D9C135 CRC64;

Query Match 96.3%; Score 79; DB 1; Length 239;
Best Local Similarity 93.3%; Pred. No. 0.85;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15

```

```

Db 158 EEEEDDEDEDD 172
|||||||:|
RESULT 6
Q8T253 PRELIMINARY; PRT; 2246 AA.
AC Q8T253;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachbat J., Dear P., Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC115683; AAL92693.2; -.
DR InterPro; IPR002110; ANK.
DR SMART; SM00248; ANK; 5.
KW Hypothetical protein.
SQ SEQUENCE 2246 AA; 265604 MW; E99C27A2E19FC67A CRC64;

Query Match 92.7%; Score 76; DB 2; Length 2246;
Best Local Similarity 86.7%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
Db 330 EEEEDDEDEDD 344
|||||||:|

RESULT 7
Q39892 PRELIMINARY; PRT; 358 AA.
AC Q39892;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nucleosome assembly protein 1.
GN Name=SNAP-1;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=Williams; TISSUE=Hypocotyl;
RX MEDLINE=96133687; PubMed=8544812;
RA Yoon H.W., Kim M.C., Lee S.Y., Hwang I., Bahk J.D., Hong J.C., Ishimi Y., Cho M.J.;
RA "Molecular cloning and functional characterization of a cDNA encoding Nucleosome assembly protein 1 (NAP-1) from soybean.";
RL Mol. Gen. Genet. 249:465-473(1995).
CC -!- SIMILARITY: Belongs to the nucleosome assembly protein (NAP) family.
CC EMBL; L38856; AAA88792.1; -.
DR PIR; S60892; S60892.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

```

```
DR GO: 0006334; P:nucleosome assembly; IEA.
DR InterPro; IPR002164; NAP_family.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00956; NAP; 1.
SQ SEQUENCE 358 AA; 40932 MW; C71F9AE9BA04699D CRC64;
Query Match 91.5%; Score 75; DB 2; Length 358;
Best Local Similarity 86.7%; Pred. No. 3.1;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDEDD 15
:||||:|||||
Db 307 DEEDEDDEDEDD 321

RESULT 8
ID Q7RKZ2 PRELIMINARY; PRT; 1260 AA.
AC Q7RKZ2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY02756;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiolini S.V., Suh B.B., Kooij T.W., Perlea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.J., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000765; EAA22241.1; -.
DR InterPro; IPR008941; TPR-like.
KW Hypothetical protein.
SQ SEQUENCE 1260 AA; 146769 MW; 84BE885A00CF5767 CRC64;
Query Match 91.5%; Score 75; DB 2; Length 1260;
Best Local Similarity 86.7%; Pred. No. 10;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDEDD 15
|||||:|||||
Db 127 EEEEDDDDEDEDE 141

RESULT 9
ID NUC1_CHICK STANDARD; PRT; 694 AA.
AC P15771;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Nucleolin (Protein C23).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
```

```
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90206792; PubMed=2320420;
RA Maridor G., Nigg E.A.;
RT "cDNA sequences of chicken nucleolin/C23 and N038/B23, two major
RT nucleolar proteins.";
RL Nucleic Acids Res. 18:1286-1286(1990).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=90304215; PubMed=2114180; DOI=10.1016/0167-4781(90)90032-W;
RA Maridor G., Krek W., Nigg E.A.;
RT "Structure and developmental expression of chicken nucleolin and N038:
RT coordinate expression of two abundant non-ribosomal nucleolar
RT proteins.";
RL Biochim. Biophys. Acta 1049:126-133(1990).
RN [3]
RP SEQUENCE OF 407-694 FROM N.A.
RX MEDLINE=89119560; PubMed=2914325; DOI=10.1016/0092-8674(89)90241-9;
RA Borer R.A., Lehner C.F., Eppenberger H.M., Nigg E.A.;
RT "Major nucleolar proteins shuttle between nucleus and cytoplasm.";
RL Cell 56:379-390(1989).
CC -!- FUNCTION: Nucleolin is the major nucleolar protein of growing
CC eukaryotic cells. It is found associated with intranucleolar
CC chromatin and preribosomal particles. It induces chromatin
CC decondensation by binding to histone H1. It is thought to play a
CC role in pre-rRNA transcription and ribosome assembly.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC
CC EMBL; X17199; CAA35080.1; -.
DR EMBL; M21791; AAA48983.1; -.
DR FIR; S08414; DNCNHL.
DR HSSP; P08199; 1FJC.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; RRM_1; 4.
DR SMART; SM00360; RRM; 4.
DR PROSITE; P850102; RRM; 4.
DR DNA-binding; Methylation; Nuclear protein; Phosphorylation; Repeat;
KW RNA-binding.
FT DOMAIN 124 141 Asp/Glu-rich (acidic).
FT DOMAIN 170 192 Asp/Glu-rich (acidic).
FT DOMAIN 217 247 Asp/Glu-rich (acidic).
FT DOMAIN 281 357 RNA-binding (RRM) 1.
FT DOMAIN 371 445 RNA-binding (RRM) 2.
FT DOMAIN 461 535 RNA-binding (RRM) 3.
FT DOMAIN 553 628 RNA-binding (RRM) 4.
FT DOMAIN 630 682 Arg/Gly/Phe-rich.
FT DOMAIN 55 90 5 X 7 AA tandem repeats of X-T-P-X-K-K-X.
FT REPEAT 55 61 1.
FT REPEAT 62 68 2.
FT REPEAT 69 75 3.
FT REPEAT 76 82 4.
FT REPEAT 84 90 5.
FT MOD_RES 116 136 Phosphoserine (By similarity).
FT MOD_RES 136 136 Phosphoserine (By similarity).
FT MOD_RES 171 171 Phosphoserine (By similarity).
FT CONFLICT 419 419 A -> R (in Ref. 3).
FT CONFLICT 520 520 N -> T (in Ref. 3).
SQ SEQUENCE 694 AA; 75640 MW; 7996C504BE9459A1 CRC64;
Query Match 90.2%; Score 74; DB 1; Length 694;
Best Local Similarity 86.7%; Pred. No. 7.2;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1 EEEEDDDDEDEDD 15
Db 173 EEEEDDDDEDEDD 187

RESULT 10
Q8BWG3 PRELIMINARY; PRT; 370 AA.
AC Q8BWG3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
DE library, clone: D63000SH08 product: hypothetical Zinc finger, C2H2 type
DE containing protein, full insert sequence.
GN Name=BC035954;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama Y., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizune W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,

```

```

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK052616; BAC35064.1; -
DR MGD; MGI:2673697; BC035954.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR SMART; SM00355; ZNF_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 370 AA; 40151 MW; EAE899898FD17706 CRC64;

Query Match 89.0%; Score 73; DB 2; Length 370;
Best Local Similarity 86.7%; Pred. No. 5;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEEEDDDDEDEDD 15
Db 306 EEAEDDDDEDEDD 320

RESULT 11
Q8CI82 PRELIMINARY; PRT; 409 AA.
AC Q8CI82;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BC035954 protein.
GN Name=BC035954;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035954; AAH35954.1; -
DR MGD; MGI:2673697; BC035954.
DR GO; GO:0005634; C:nucleus; IEA.

```

DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro: IPR007087; ZnF_C2H2; 3.
 DR SMART; SM00355; ZnF_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
 SQ SEQUENCE 409 AA; 43951 MW; 60EB02611DA9EC5 CRC64;

Query Match 89.0%; Score 73; DB 2; Length 409;
 Best Local Similarity 86.7%; Pred. No. 5.5;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||:|:|:|
 Db 345 EEEDDDDDDEDD 359

RESULT 12

Q92243 ID Q92243 PRELIMINARY; PRT; 457 AA.
 AC Q92243;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE NNK3.
 GN Name=C80913;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99097352; PubMed=9878255; DOI=10.1006/geno.1998.5609;
 RA Van Leuven F., Torrekens S., Moehrs D., Hilliker C., Buellens M.,
 RA Bollen M., Delabie J.;
 RT "Molecular cloning of a gene on chromosome 19q12 coding for a novel
 RT intracellular protein: analysis of expression in human and mouse
 RT tissues and in human tumor cells, particularly Reed-Sternberg cells in
 RT Hodgkin disease.";
 RL Genomics 54:511-520(1998).
 DR EMBL; AF091096; AAD08680.1; -.
 DR MGD; MGI:1342294; C80913.
 DR GO: GO:0016272; C:prefoldin complex; IEA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR GO: GO:0006457; P:protein folding; IEA.
 DR InterPro; IPR004127; PfD_alpha-like.
 DR InterPro; IPR009053; Prefoldin.
 DR Pfam; PF02996; Prefoldin; 1.
 SQ SEQUENCE 457 AA; 50647 MW; 1A07C4C4B89E13BA CRC64;

Query Match 89.0%; Score 73; DB 2; Length 457;
 Best Local Similarity 80.0%; Pred. No. 6.1;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||:|:|:|
 Db 229 EEEDDDDDDEDD 243

RESULT 13

Q6GNP2 ID Q6GNP2 PRELIMINARY; PRT; 625 AA.
 AC Q6GNP2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MGC80973 protein.
 GN Name=MGC80973;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;

RN SEQUENCE FROM N.A.
 RP TISSUE=Embryo;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/advy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC073462; AAH73462.1; -.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003676; F:nucleic acid binding; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR009079; 4 helix cytokine.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR Pfam; PF00096; zf-C2H2; 9.
 DR ProDom; PD000003; ZnF_C2H2; 1.
 DR SMART; SM00355; ZnF_C2H2; 9.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 6.
 SQ SEQUENCE 625 AA; 71186 MW; 9AC1B6C726B46E7D CRC64;

Query Match 89.0%; Score 73; DB 2; Length 625;
 Best Local Similarity 80.0%; Pred. No. 8.2;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
 |||||:|:|:|
 Db 104 EEEDDDDDDEDD 118

RESULT 14

ID IFH1 YEAST STANDARD; PRT; 1085 AA.
 AC P39520;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE IFH1 protein (RRP3 protein).
 GN Name=IFH1; Synonyms=RRP3; OrderedLocusNames=YLR223C; ORFNames=L8083.9;
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;


```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=ATCC 28383 / FL100;
RX  MEDLINE=95304839; PubMed=7785326;
RA  Cherel I., Thuriaux P.;
RT  "The IFH1 gene product interacts with a fork head protein in
RL  Saccharomyces cerevisiae.";
RL  Yeast 11:261-270(1995).
[2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=S288c / AB972;
RX  MEDLINE=973113267; PubMed=9169871;
RA  Johnson M., Hillier L., Riles L., Albermann K., Andre B., Ansoerge W.,
RA  Benes V., Brueckner M., Delius H., Dubois E., Dueterhoeft A.,
RA  Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA  Heuss-Netzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA  Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA  Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA  Portetelle D., Furnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA  Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA  Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA  Vierdeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA  Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT  "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL  Nature 387:87-90(1997).
CC  -1- FUNCTION: Controls the pre-rRNA processing machinery in
CC  conjunction with PHL1. Could convert PHL1 from a repressor to an
CC  activator.
CC  -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; Z29488; CAA82624.1; -.
DR  EMBL; U19027; AAB67412.1; -.
DR  PIR; S55352; S55352.
DR  InAct; F39520; -.
DR  Germonline; 142285; -.
DR  SGD; S000004213; IFH1.
DR  GO; GO:0005634; C:nucleus; IDA.
DR  GO; GO:0006348; P:chromatin silencing at telomere; IMP.
DR  GO; GO:0006364; P:rRNA processing; IGI.
DR  GO; GO:0006364; P:rRNA processing; IGI.
KW  Nuclear protein; Transcription regulation.
FT  DOMAIN 122 163 Asp/Glu-rich (highly acidic).
SQ  SEQUENCE 1085 AA; 122491 MW; BE1C7DEF06213FE0 CRC64;

Query Match 89.0%; Score 73; DB 1; Length 1085;
Best Local Similarity 80.0%; Pred.No. 14;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDDDEDEDD 15
Db 133 EEEEDDDDEDDDD 147

RESULT 15
Q7S403
ID Q7S403 PRELIMINARY; PRT; 1359 AA.
AC Q7S403;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU05902.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
```

```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=OR74A;
RA  Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA  Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Renman B.,
RA  Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA  Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA  Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA  Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA  Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA  Kamal M., Kamvysselis M., Mauceli E., Bielek C., Rudd S., Frisman D.,
RA  Kryzofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA  Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA  DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA  Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA  Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA  Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT  "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL  Nature 0:0-0(2003).
CC  -1- SIMILARITY: Contains 2 WD repeats.
CC  -1- CAUTION: The sequence shown here is derived from an
CC  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC  preliminary data.
DR  EMBL; AABX01000384; EAA30224.1; -.
DR  InterPro; IPR008938; ARM.
DR  InterPro; IPR000547; Clathrin_repeat.
DR  InterPro; IPR001199; Cyt_B5.
DR  InterPro; IPR001680; WD40.
DR  InterPro; IPR011046; WD40_like.
DR  Pfam; PF00637; Clathrin; 1.
DR  Pfam; PF00400; WD40; 2.
DR  PROSITE; PS00191; CYTOCHROME B5_1; UNKNOWN_1.
KW  Hypothetical protein; Repeat; WD repeat.
SQ  SEQUENCE 1359 AA; 149201 MW; BDCB61B70764D641 CRC64;

Query Match 89.0%; Score 73; DB 2; Length 1359;
Best Local Similarity 86.7%; Pred.No. 17;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEEEDDDDEDEDD 15
Db 76 EETEEDDEDEDEDD 90

Search completed: September 20, 2005, 12:41:43
Job time : 9.61934 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 1.68911 Seconds
(without alignments)
854.447 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82

Sequence: 1 EEEDEDEDEDD 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	91.5	358	2 S60892	nucleosome assembl
2	74	90.2	694	1 DNCNHL	nucleolin - chicke
3	73	89.0	1085	2 S55352	IFH1 protein - yea
4	72	87.8	178	2 T51159	HMG protein (impor
5	72	87.8	288	2 T22846	hypothetical prote
6	71	86.6	667	2 JC5889	OS-9 protein precu
7	71	86.6	712	2 JH0148	nucleolin - rat
8	71	86.6	727	2 JC5113	ribosomal transcri
9	71	86.6	764	2 JC5112	ribosomal transcri
10	70	85.4	215	2 S02826	nonhistone chromos
11	70	85.4	562	2 S38149	SIS2 protein - yea
12	70	85.4	707	2 A35804	nucleolin - human
13	70	85.4	802	1 S48529	NAB3 protein - yea
14	70	85.4	1105	2 T18295	Ap-3 adaptor compl
15	69	84.1	279	2 S30766	ASF1 protein - yea
16	69	84.1	352	2 T51687	neurogenic differe
17	69	84.1	636	2 T51893	related to Che-1 p
18	69	84.1	867	2 T40336	probable MGF trans
19	69	84.1	2182	2 T28634	variant-specific s
20	68	82.9	379	2 C84577	probable nucleosom
21	68	82.9	389	2 T27574	hypothetical prote
22	68	82.9	565	2 T07424	phosphonositide-s
23	68	82.9	678	2 A54514	glutamic acid-rich
24	68	82.9	700	2 S38426	chaperonin 60 prec
25	68	82.9	727	2 S18193	autoantigen NOR-90
26	68	82.9	764	2 S09318	transcription fact
27	68	82.9	792	2 T42963	hypothetical prote
28	68	82.9	806	2 T24169	hypothetical prote
29	68	82.9	906	2 T24166	hypothetical prote

probable ubiquitin
dopamine- and cAMP
nonhistone chromos
HMG-1 - chicken
nonhistone chromos
non-histone chromo
31K proliferation
nucleophosmin - Af
probable histone d
protein T2737.4 [i
hypothetical prote
transcription fact
transcription fact
hypothetical prote
probable translati

ALIGNMENTS

RESULT 1

S60892

nucleosome assembly protein 1 - soybean

C;Species: Glycine max (soybean)

C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C;Accession: S60892

R;Yoon, H.W.; Kim, M.C.; Lee, S.Y.; Hwang, I.; Bahk, J.D.; Hong, J.C.; Ishimi, Y.; Cho, I.

Mol. Gen. Genet. 249: 465-473, 1995

A;Title: Molecular cloning and functional characterization of a cDNA encoding nucleosome

A;Reference number: S60892; MUID:96133687; PMID:8544812

A;Accession: S60892

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-358 <YOO>

A;Cross-references: UNIPROT:Q39892; EMBL:L38856; NID:g1161251; PIDN:AAA88792.1; PID:g1161

Query Match 91.5%; Score 75; DB 2; Length 358;

Best Local Similarity 86.7%; Pred. No. 0.28;

Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEEDEDEDEDD 15

Db 307 DEEDEDDEDD 321

RESULT 2

DNCNHL

nucleolin - chicken

N;Alternate names: nucleolar protein C23

C;Species: Gallus gallus (chicken)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C;Accession: S08414; S10766; A32725; I50397; B30099

R;Maridor, G.; Nigg, E.A.

Nucleic Acids Res. 18, 1286, 1990

A;Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar protei

A;Reference number: S08414; MUID:90206792; PMID:2320420

A;Accession: S08414

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-694 <NAR>

A;Cross-references: UNIPROT:P15771; EMBL:X1719; NID:g63710; PIDN:CAA35060.1; PID:g63711

R;Maridor, G.; Krek, W.; Nigg, E.A.

Biochim. Biophys. Acta 1049, 126-133, 1990

A;Title: Structure and developmental expression of chicken nucleolin and NO38: coordinate

A;Reference number: S10766; MUID:90304215; PMID:2114180

A;Accession: S10766

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-694 <NAR>

R;Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.

Cell 60, 791-801, 1990

A;Title: Identification of major nucleolar proteins as candidate mitotic substrates of c

A;Reference number: A32725; MUID:90182668; PMID:2178776
A;Accession: A32725
A;Molecule type: protein
A;Residues: 56-62;63-109;197-214 <PET>
R;Borer, R.A.; Lehner, C.F.; Epenberger, H.M.; Nigg, E.A.
Cell 56, 379-390, 1989
A;Title: Major Nuclear Proteins Shuttle between Nucleus and Cytoplasm.
A;Reference number: I50397; MUID:89119560; PMID:2914325
A;Accession: I50397
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 407-418,'R',420-519,'T',521-694 <BOR>
A;Cross-references: GB:M21791; NID:G212411; PIDN:AAA48983.1; PID:G212412
C;Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorganization of nucleolin; ribonucleoprotein repeat homology
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; duplication; nucleus; phosphoprotein; RNA binding; RNA binding; F;1-247/Domain: acidic <ACI>
F;1-247/Domain: acidic <ACI>
F;54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])
F;254-262/Region: nuclear location signal
F;267-275/Region: nuclear location signal
F;282-347/Domain: ribonucleoprotein repeat homology <RRM1>
F;283-361/Domain: RNA binding #status predicted <RNA1>
F;372-435/Domain: ribonucleoprotein repeat homology <RRM2>
F;373-448/Domain: RNA binding #status predicted <RNA2>
F;462-525/Domain: ribonucleoprotein repeat homology <RRM3>
F;463-538/Domain: RNA binding #status predicted <RNA3>
F;554-618/Domain: ribonucleoprotein repeat homology <RRM4>
F;555-631/Domain: RNA binding #status predicted <RNA4>
F;632-694/Domain: glycine/arginine-rich <GRR>
F;55,63,70,77,85/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predicted

Query Match 90.2%; Score 74; DB 1; Length 694;
Best Local Similarity 86.7%; Pred.No. 0.68;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDED 15
Db 173 EEEEEDEDED 187

RESULT 3
S55352
NFH1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein I8083.9; protein YLR223c; RRP3 protein
C;Species: Saccharomyces cerevisiae
C;Date: 10-Oct-1995 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55352; S51446; S47477
R;Cherel, I.; Thuriaux, P.
Yeast 11, 261-270, 1995
A;Title: The IFH1 gene product interacts with a fork head protein in Saccharomyces cerevisiae
A;Reference number: S55352; MUID:95304839; PMID:7785326
A;Accession: S55352
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1085 <HE>
A;Cross-references: UNIPROT:P39520; EMBL:Z29488; NID:G531491; PIDN:CAA82624.1; PID:G531491; Hallsworth, K.
A;Description: The sequence of S. cerevisiae cosmid 8083.
A;Reference number: S51443
A;Accession: S51446
A;Molecule type: DNA
A;Residues: 1-1085 <HAL>
A;Cross-references: EMBL:U19027; NID:G609363; PID:G609372; MIPS:YLR223c
C;Genetics:
A;Gene: SGD:IFH1; RRP3
A;Cross-references: SGD:S0004213; MIPS:YLR223c
A;Map position: 12R

Query Match 89.0%; Score 73; DB 2; Length 1085;
Best Local Similarity 80.0%; Pred.No. 1.3;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
JCS989
OS-9 protein precursor - human
N:Contains: OS-9 protein, splice form 1; OS-9 protein, splice form 2; OS-9 protein, splice form 3
C:Species: Homo sapiens (man)
C:Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 09-Jul-2004
C:Accession: JCS989; JEO106; JEO107; JEO108
R:Kimura, Y.; Nakazawa, M.; Tsuchiya, N.; Asakawa, S.; Shimizu, N.; Yamada, M.
J. Biochem. 122, 1190-1195, 1997
A:Title: Genomic organization of the OS-9 gene amplified in human sarcomas.
A:Reference number: JCS989; MUID:98158329; PMID:9498564
A:Accession: JCS989
A:Molecule type: DNA
A:Residues: 1-667 <KIM1>
A:Cross-references: UNIPROT:Q13438; DDBJ:AB002806; NID:G2780782; PIDN:BAA24363.1; PID:d1025275
A:Experimental source: sarcomas
A>Note: neither the complete nucleic acid sequence nor the complete translation are shown
R:Kimura, Y.; Nakazawa, M.; Yamada, M.
J. Biochem. 123, 876-882, 1998
A:Title: Cloning and characterization of three isoforms of OS-9 cDNA and expression of the OS-9 protein
A:Reference number: JEO108; MUID:98230694; PMID:9562620
A:Accession: JEO106
A:Molecule type: mRNA
A:Residues: 1-667 <KIM2>
A:Cross-references: DDBJ:AB002806; NID:G2780782; PIDN:BAA24363.1; PID:d1025275; PID:G2780782
A:Accession: JEO107
A:Molecule type: mRNA
A:Residues: 1-534, 590-667 <KIM3>
A:Cross-references: DDBJ:AB002806
A:Accession: JEO108
A:Molecule type: mRNA
A:Residues: 1-455, 471-534, 590-667 <KIM4>
A:Cross-references: DDBJ:AB002806
C:Comment: This protein is involved in amplification and overexpression of various tumors
C:Genetics:
A:Gene: OS-9
A:Cross-references: GDB:9958646
A:Map position: 12q13-12q15
C:Keywords: alternative splicing; carcinogenesis; glycoprotein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:16-667/Product: OS-9 protein, splice form 1 #status predicted <MAT1>
F:26-534, 590-667/Product: OS-9 protein, splice form 2 #status predicted <MAT2>
F:26-455, 471-534, 590-667/Product: OS-9 protein, splice form 3 #status predicted <MAT3>
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 86.6%; Score 71; DB 2; Length 667;
Best Local Similarity 80.0%; Pred. No. 1.3;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
|||||:|||||:
Db 414 EEEEDDEDEDED 428

RESULT 7
JH0148
nucleolin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C:Accession: JH0148; A24088; I63130
R:Bourbon, H.M.; Amalric, F.
Gene 88, 187-196, 1990
A:Title: Nucleolin gene organization in rodents: highly conserved sequences within three exons
A:Reference number: JH0148; MUID:90269607; PMID:2347493
A:Accession: JH0148
A:Molecule type: DNA
A:Residues: 1-712 <BOU>
A>Note: the authors translated the initiation codon GTG for residue 1 as Val
R:Lischwe, M.A.; Cook, R.G.; Ahn, Y.S.; Yeoman, L.C.; Busch, H.
Biochemistry 24, 6025-6028, 1985
A:Title: Clustering of glycine and Ng,Ng-dimethylarginine in nucleolar protein C23.
A:Reference number: A24088; MUID:86104094; PMID:4084504

A:Accession: A24088
A:Molecule type: protein
A:Residues: 651-703 <LIS>
R:Bourbon, H.
Gene 89, 73-84, 1988
A:Title: Sequence and structure of the nucleolin promoter in rodents: Characterization of the nucleolin promoter in rodents: Characterization of the nucleolin promoter in rodents
A:Reference number: I48118; MUID:89121496; PMID:2906027
A:Accession: I63130
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 'MV', 2-44 <RES>
A:Cross-references: GB:M2090; NID:G205793; PIDN:AAA1733.1; PID:G205794
C:Comment: This protein is the major nucleolar-specific protein in eukaryotic exponential phase cells
C:Genetics:
A:Gene: nuc
A:Start codon: GTG
A:Introns: 5/3; 44/3; 211/1; 274/1; 303/1; 350/2; 392/1; 433/2; 485/1; 526/2; 571/1; 613/1
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; nucleus
F:311-376/Domain: ribonucleoprotein repeat homology <RRM1>
F:397-459/Domain: ribonucleoprotein repeat homology <RRM2>
F:489-552/Domain: ribonucleoprotein repeat homology <RRM3>
F:575-639/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 86.6%; Score 71; DB 2; Length 712;
Best Local Similarity 80.0%; Pred. No. 1.4;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
|||||:|||||:
Db 241 EEEEDDEDEDED 255

RESULT 8
JCS113
ribosomal transcription factor UBF2 - Chinese hamster
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Aug-2004
C:Accession: JCS113
R:Bollivar, J.; Goenechea, L.G.; Grenett, H.; Pendon, C.; Valdivia, M.M.
Gene 176, 257-258, 1996
A:Title: Cloning and sequencing of the genes encoding the hamster ribosomal transcription factor UBF2
A:Reference number: JCS112; MUID:97075939; PMID:8918262
A:Accession: JCS113
A:Molecule type: mRNA
A:Residues: 1-727 <BOL>
A:Cross-references: UNIPROT:Q60460; GB:L42571
C:Comment: This factor binds to the ribosomal RNA gene's promoter and forms a stable pre-ribosomal transcription factor UBF2 - Chinese hamster
C:Superfamily: HMG box homology
F:109-184/Domain: HMG box homology <HMG1>
F:257-329/Domain: HMG box homology <HMG3>
F:445-516/Domain: HMG box homology <HMG5>

Query Match 86.6%; Score 71; DB 2; Length 727;
Best Local Similarity 80.0%; Pred. No. 1.4;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
|||||:|||||:
Db 646 EEEEDDEDEDED 660

RESULT 9
JCS112
ribosomal transcription factor UBF1 - Chinese hamster
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Aug-2004
C:Accession: JCS112
R:Bollivar, J.; Goenechea, L.G.; Grenett, H.; Pendon, C.; Valdivia, M.M.
Gene 176, 257-258, 1996
A:Title: Cloning and sequencing of the genes encoding the hamster ribosomal transcription factor UBF1
A:Reference number: JCS112; MUID:97075939; PMID:8918262
A:Accession: JCS112

A;Molecule type: mRNA
A;Residues: 1-764 <BOL>
A;Cross-references: UNIPROT:Q60459; GB:I42570
C;Comment: This factor binds to the ribosomal RNA gene's promoter and forms a stable pre-
C;Superfamily: HMG box homology
F;109-184/Domain: HMG box homology <HMG1>
F;193-268/Domain: HMG box homology <HMG2>
F;294-366/Domain: HMG box homology <HMG3>

Query Match 86.4%; Score 71; DB 2; Length 764;
Best Local Similarity 80.0%; Pred. No. 1.5;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
|||:||||:|:||||
Db 683 EEEEDDEDEDD 697

RESULT 10
S02826
nonhistone chromosomal protein HMG-1 - human
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S02826; A33178; G33178
R;Wen, L.; Huang, J.K.; Johnson, B.H.; Reeck, G.R.
Nucleic Acids Res. 17, 1197-1214, 1989
A;Title: A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-1.
A;Reference number: S02826; MUID:89160247; PMID:2922262
A;Accession: S02826
A;Molecule type: mRNA
A;Residues: 1-215 <WEN>
A;Cross-references: UNIPROT:P09429; EMBL:X12597; NID:G32326; PIDN:CAA31110.1; PID:G32327
R;Ward, L.D.; Hong, J.; Whitehead, R.H.; Simpson, R.J.
Electrophoresis 11, 883-891, 1990
A;Title: Development of a database of amino acid sequences for human colon carcinoma pro
A;Reference number: A33178; MUID:91176935; PMID:2079031
A;Accession: A33178
A;Molecule type: protein
A;Residues: 2-13, 'XXF' <WAR>
A;Accession: G33178
A;Molecule type: protein
A;Residues: 2-13, 'XX', 16-22 <WA2>
C;Genetics:
A;Gene: GDB:HWG1
A;Cross-references: GDB:133789; OMIM:163905
A;Map position: 13q12-13q12
C;Superfamily: nonhistone chromosomal protein HMG-2; HMG box homology
C;Keywords: chromosomal protein; DNA binding; nucleus
F;6-83/Domain: HMG box homology <HMG1>
F;92-166/Domain: HMG box homology <HMG2>

Query Match 85.4%; Score 70; DB 2; Length 215;
Best Local Similarity 80.0%; Pred. No. 0.56;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15
|||:||||:|:||||
Db 197 EEEEDDEDEDD 211

RESULT 11
S38149
SIS2 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YKR072C
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: S38149; S54982; S43078
R;Pohl, T.M.; Pohl, F.M.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37897
A;Accession: S38149
A;Molecule type: DNA
A;Residues: 1-562 <POH>

A;Cross-references: UNIPROT:P36024; EMBL:Z28297; NID:G486544; PID:G486545; MIPS:YKR072C
A;Experimental source: strain S288C
R;di Como, C.J.; Bose, R.; Arndt, K.T.
Genetics 139, 95-107, 1995
A;Title: Overexpression of SIS2, which contains an extremely acidic region, increases the
A;Reference number: S54982; MUID:95220693; PMID:7705654
A;Accession: S54982
A;Molecule type: DNA
A;Residues: 1-562 <DIW>
A;Cross-references: EMBL:U01878; NID:G430983; PIDN:AAA80000.1; PID:G430984
C;Genetics:
A;Gene: SGD:SIS2
A;Cross-references: SGD:S0001780; MIPS:YKR072C
A;Map position: 11R
C;Keywords: nucleus

Query Match 85.4%; Score 70; DB 2; Length 562;
Best Local Similarity 85.7%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEEEDDEDEDD 15
|||||:||||:|
Db 496 EEEEDDEDEDD 509

RESULT 12
A35804
nucleolin - human
N;Alternate names: phosphoprotein ppl00; protein B50; protein C23
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text_change 09-Jul-2004
C;Accession: A35804; S04631; A48138; A55996
R;Srivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
J. Biol. Chem. 265, 14922-14931, 1990
A;Title: Genomic organization and chromosomal localization of the human nucleolin gene.
A;Reference number: A35804; MUID:90368666; PMID:2394707
A;Accession: A35804
A;Molecule type: DNA
A;Residues: 1-707 <SR1>
A;Cross-references: UNIPROT:P19338; GB:M60858; GB:J05584; NID:G189305; PIDN:AAA59954.1; I
R;Srivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
FEBS Lett. 250, 99-105, 1989
A;Title: Cloning and sequencing of the human nucleolin cDNA.
A;Reference number: S04631; MUID:89290043; PMID:2737305
A;Accession: S04631
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-707 <SR2>
R;Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.
Mol. Cell. Biol. 13, 4301-4310, 1993
A;Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and t
A;Reference number: A48138; MUID:93309464; PMID:8321232
A;Accession: A48138
A;Molecule type: protein
A;Residues: 458-474 <ISH>
A;Experimental source: HeLa cell nuclei
A;Note: sequence extracted from NCBI backbone (NCBIP:134645)
R;Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.
Biochemistry 33, 14696-14706, 1994
A;Title: Major cell surface-located protein substrates of an ecto-protein kinase are hom
A;Reference number: A55996; MUID:95086063; PMID:7993898
A;Accession: A55996
A;Molecule type: protein
A;Residues: 231-236;349-362;399-403;458-461;655-656, 'X', 658-660 <JOR>
C;Genetics:
A;Gene: GDB:NCL
A;Cross-references: GDB:125908; OMIM:164035
A;Map position: 2q12-2qter
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleolus; nucleus; phosphoprotein; RNA binding
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>

F;487-550/Domain: ribonucleoprotein repeat homology <RRM3>
F;573-634/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 85.4%; Score 70; DB 2; Length 707;
Best Local Similarity 80.0%; Pred. No. 1.8;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
:|||||:|||||
Db 146 DEEDDDSEDEDD 160

RESULT 13

S48529
NAB3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: probable RNA/asDNA-binding protein HMD1; protein P1945; protein YPL19
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S48529; S65209; S60122
R:Wilson, S.M.; Oberdorf, A.M.; Datar, K.V.; Swedlow, J.R.; Paddy, M.R.; Swanson, M.S.
A:Description: Characterization of Nuclear Polyadenylated RNA-Binding Proteins from Sac
submitted to the EMBL Data Library, January 1994
A:Reference number: S48529
A:Accession: S48529
A:Molecule type: DNA
A:Residues: 1-802 <WIL>
A:Cross-references: UNIPROT:P38996; EMBL:U05314; NID:g476219; PID:g476220
R:Rieger, M.; Mueller-Auer, S.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S65202
A:Accession: S65209
A:Molecule type: DNA
A:Residues: 1-802 <RIE>
A:Cross-references: EMBL:Z73546; NID:g1370396; PID:g1370397; GSPDB:GN00016; MIPS:YPL190
A:Experimental source: strain S288C (AB372)
R:Sugimoto, K.; Matsumoto, K.; Kornberg, R.D.; Reed, S.I.; Wittenberg, C.
Mol. Gen. Genet. 248, 712-718, 1995
A:Title: Dosage suppressors of the dominant G1 cyclin mutant CLN3-2: identification of a
A:Reference number: S60122; MUID:96069710; PMID:7476874
A:Accession: S60122
A>Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-340,'I',342-802 <SUG>
A:Cross-references: GB:D37935; NID:g1235749; PID:g1235750
C:Genetics:
A:Gene: SGD:NAB3; NAB3; HMD1; MIPS:YPL190C
A:Cross-references: SGD:S000611
A:Map position: 16L
C:Superfamily: yeast NAB3 protein; ribonucleoprotein repeat homology
C:Keywords: nucleus; RNA binding
F;331-396/Domain: ribonucleoprotein repeat homology <RRM1>

Query Match 85.4%; Score 70; DB 1; Length 802;
Best Local Similarity 73.3%; Pred. No. 2;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
:|||||:|||||
Db 112 EEEEDDDDDDDDD 126

RESULT 14

T18295
Ap-3 adaptor complex beta3a chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18295
R:Feng, L.; Seymour, A.B.; Jiang, S.Y.; To, A.; Peden, A.A.; Novak, E.K.; Zhen, L.; Rusi
Hum. Mol. Genet. 8, 323-330, 1999
A:Title: The beta3a subunit gene (Ap3b1) of the AP-3 adaptor complex is altered in the m
A:Reference number: Z18864; MUID:99135912; PMID:9931340
A:Accession: T18295
A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA
A:Residues: 1-1105 <FEN>
A:Cross-references: UNIPROT:O9Z1T1; EMBL:AF103809; NID:g3885987; PID:g3885988; PIDN:AAC7
A:Experimental source: strain C3H/HeJ
C:Genetics:
A:Gene: Ap3b1
A:Map position: 13

Query Match 85.4%; Score 70; DB 2; Length 1105;
Best Local Similarity 80.0%; Pred. No. 2.7;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
:|||||:|||||
Db 681 EEEEDDEDEDEEE 695

RESULT 15

S30766
ASF1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J0755; protein YJL115W
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C:Accession: S30766; S56896
R:Le, S.; Sternglanz, R.
submitted to the EMBL Data Library, December 1992
A:Reference number: S30766
A:Accession: S30766
A:Molecule type: DNA
A:Residues: 1-279 <LES>
A:Cross-references: UNIPROT:P32447; EMBL:L07593; NID:g171090; PIDN:AAC37512.1; PID:g1710
R:Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56891
A:Accession: S56896
A:Molecule type: DNA
A:Residues: 1-279 <CZI>
A:Cross-references: EMBL:Z49390; NID:g1008303; PIDN:CAAB9410.1; PID:g1008304; MIPS:YJL11
C:Genetics:
A:Gene: SGD:ASF1
A:Cross-references: SGD:S0003651; MIPS:YJL115W
A:Map position: 10L

Query Match 84.1%; Score 69; DB 2; Length 279;
Best Local Similarity 73.3%; Pred. No. 0.91;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
:|||||:|||||
Db 178 EEEEDDEDEDEDD 192

Search completed: September 20, 2005, 12:43:19
Job time : 2.68911 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 15.4663 Seconds
(without alignments)
650.171 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEDDEDEDEDDVSEGSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	26	4	AAB82810 Human low
2	136	100.0	217	2	AAW49041 Human low
3	136	100.0	217	4	AAB82803 Human low
4	136	100.0	241	3	AAB58957 Breast an
5	136	100.0	538	4	AAB82806 Human low
6	104.5	76.8	28	4	AAB82816 Rabbit lo
7	104.5	76.8	232	4	AAB82799 Rabbit lo
8	104.5	76.8	252	4	AAB82800 Rabbit lo
9	104.5	76.8	317	2	AAW49038 Rabbit lo
10	104.5	76.8	317	4	AAB82798 Rabbit lo
11	104.5	76.8	550	4	AAB82807 Rabbit lo
12	88	64.7	162	2	AAR30641 pCENP-B-1
13	88	64.7	183	3	AAB43785 Human can
14	88	64.7	183	4	AAG74067 Human col
15	88	64.7	487	8	ABM82201 Tumour-as
16	88	64.7	594	2	AAR34936 CENP-B.7
17	88	64.7	594	8	ADJ26909 Human cen
18	88	64.7	599	7	ADP65135 Human cen
19	88	64.7	599	7	ADP65197 Human cen
20	83.5	61.4	155	3	AAG17704 Arabidops
21	83.5	61.4	305	3	AAG26185 Arabidops
22	83.5	61.4	305	4	AAB67814 Amino aci
23	83.5	61.4	306	5	ABM93572 Herbicida
24	83	61.0	152	3	AAB07863 Amino aci
25	83	61.0	427	5	ABP73549 Candida a

26	83	61.0	539	3	AAV53050	Human sec
27	83	61.0	1501	6	AAE36115	Human chr
28	83	61.0	1953	7	ADJ80153	Novel hum
29	82	60.3	15	4	AAB82809	Human low
30	82	60.3	343	4	ABB58472	Drosophil
31	82	60.3	358	6	ABB84611	Soybean n
32	82	60.3	712	7	ADBE1623	Ades1623
33	82	60.3	712	7	ADDA6090	Rat Prote
34	82	60.3	712	7	ADE57828	Rat Prote
35	82	60.3	712	7	ADE57830	Rat Prote
36	82	60.3	759	4	ABB63935	Drosophil
37	81	59.6	444	6	ABR41713	Human DIT
38	81	59.6	768	8	ADP98855	C. albica
39	81	59.6	1161	8	ADM57329	Recombina
40	81	59.6	1162	3	AAV96255	Kaposi's
41	81	59.6	1162	3	AAV58500	HHV8 ORF
42	81	59.6	1162	4	AAB62331	Amino aci
43	81	59.6	1162	5	ABB05621	Kaposi's
44	81	59.6	1162	8	ADJ65096	HHV8 late
45	80	58.8	77	4	AAM15201	Peptide #

ALIGNMENTS

RESULT 1

AAB82810

ID AAB82810 standard; peptide; 26 AA.

XX AAB82810;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.

KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;

KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Homo sapiens.

XX WO200164874-A2.

PD 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

PS Claim 14(a); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues

CC 329-354 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of

CC the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active

CC fragments and analogues of LBP, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs.

CC Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for

CC atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in

CC structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine

CC compositions. Note: the present sequence is not shown in the
 CC specification but is derived from the human LBP-2 sequence given in
 CC figure 7A (see AAB82806)
 CC
 XX
 SQ Sequence 26 AA;

Query Match 100.0%; Score 136; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
 |||||
 Db 1 EREDDDEDEDDVSGSEVPESD 26

RESULT 2
 AAW49041
 ID AAW49041 standard; protein; 217 AA.
 XX
 AC AAW49041;
 XX
 DT 09-NOV-1998 (first entry)
 XX
 DE Human low density lipoprotein binding protein LBP-2.
 XX
 XW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
 KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 8..33
 FT /note= "Claim 2"
 FT Peptide 8..22
 FT /note= "Claim 2"
 FT Peptide 23..33
 FT /note= "Claim 2"
 FT Peptide 208..217
 FT /note= "Claim 2"
 XX
 PN WO9823282-A1.
 XX
 PD 04-JUN-1998.
 XX
 PF 26-NOV-1997; 97WO-US021857.
 XX
 PR 27-NOV-1996; 96US-0031930P.
 PR 03-JUN-1997; 97US-0048547P.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 DR WPI; 1998-322455/28.
 DR N-PSDB; AAV32838.
 XX
 FT Nucleic acid encoding low density lipoprotein binding proteins and
 FT related vectors - transformed cells, proteins, and modulators of binding,
 FT useful for treatment and diagnosis of atherosclerosis and for identifying
 FT subjects at risk.
 XX
 PS Claim 1; Fig 7; 47pp; English.
 XX
 CC This polypeptide comprises novel human low density lipoprotein (LDL)
 CC binding protein LBP-2 that is capable of binding both native and methyl
 CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
 CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human
 CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP
 CC metabolism or structure is diagnostic of a risk for atherosclerosis. The
 CC invention provides methods for determining if an animal is at risk for
 CC atherosclerosis (e.g. for prenatal screening); methods for treating
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and

CC methods for treating a cell having an abnormality in LBP structure or
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
 CC well as recombinant vectors and host cells used to produce recombinant
 CC LBP
 CC
 XX
 SQ Sequence 217 AA;

Query Match 100.0%; Score 136; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2.8e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
 |||||
 Db 8 EREDDDEDEDDVSGSEVPESD 33

RESULT 3
 AAB82803
 ID AAB82803 standard; protein; 217 AA.
 XX
 AC AAB82803;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Human low density lipoprotein binding protein 2 (LBP-2).
 XX
 XW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US006356.
 XX
 PR 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 DR WPI; 2001-565505/63.
 DR N-PSDB; AAB26494.
 XX
 PT New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX
 PS Claim 13(g); Fig 7B; 143pp; English.
 XX
 CC The present sequence is that of the N-terminal portion of novel human low
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
 CC deduced from an isolated partial cDNA clone (see AAB26494). A full-length
 CC sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP
 CC polypeptides of the invention that are capable of binding to native and
 CC methylated low density lipoproteins. Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs. Methods
 CC of determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX
 SQ Sequence 217 AA;

Query Match 100.0%; Score 136; DB 4; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2.8e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26

```

Db      8 EEEEDDEDEDDVSGSEVPESD 33
|||||
RESULT 4
AAB58957
ID AAB58957 standard; protein; 241 AA.
XX AC AAB58957;
XX DT 27-MAR-2001 (first entry)
XX DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.
XX KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
XX KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX KW antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;
XX KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
XX KW Addison's disease; allergy; autoimmune haemolytic anaemia;
XX KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
XX KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
XX KW cardiovascular disorder; wound healing; neurological disease.
XX OS Homo sapiens.
XX PN WO200055173-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005881.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-611515/58.
XX DR N-PSDB; AAF21860.
XX PT New human breast and ovarian cancer associated gene sequences and the
XX PT polypeptides encoded by these genes, useful in the prevention, treatment
XX PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
XX PT neurological diseases.
XX PS Claim 11; Page 1112; 1299pp; English.
XX CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX CC associated with breast and ovarian cancer. Included in the invention are
XX CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX CC isolation and characterisation of the DNA and protein sequences of the
XX CC invention. The breast and ovarian cancer associated DNA, protein, agonist
XX CC or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic;
XX CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
XX CC antiinflammatory; antitumor; vulnery; anticonvulsant; antibacterial;
XX CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
XX CC protein sequences are used in the diagnosis of cancer, particularly
XX CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
XX CC and agonists may also be used in the diagnosis, prevention and treatment
XX CC of immune disorders e.g. Addison's disease, allergies, autoimmune
XX CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX CC cardiovascular disorders such as myocardial ischaemias; wound healing;
XX CC neurological diseases such as cerebral anoxia and epilepsy; and
XX CC infectious diseases
XX CC Sequence 241 AA;
Query Match 100.0%; Score 136; DB 3; Length 241;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1 EEEEDDEDEDDVSGSEVPESD 26
|||||
RESULT 5
AAB82806
ID AAB82806 standard; protein; 538 AA.
XX AC AAB82806;
XX DT 12-NOV-2001 (first entry)
XX DE Human low density lipoprotein binding protein 2 (LBP-2).
XX KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
XX KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX KW Homo sapiens.
XX OS WO200164874-A2.
XX PN WO200164874-A2.
XX PD 07-SEP-2001.
XX PF 28-FEB-2001; 2001WO-US006356.
XX PR 02-MAR-2000; 2000US-00517849.
XX PR 14-JUL-2000; 2000US-00616289.
XX PA (BOST-) BOSTON HEART FOUND INC.
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX DR N-PSDB; AAB26499.
XX PT New isolated low density lipoprotein binding polypeptide for treating,
XX PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX PS Claim 13(j); Fig 7A; 143pp; English.
XX CC The present sequence is that of novel human low density lipoprotein
XX CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the
XX CC coding region of isolated genomic DNA (see AAB26499). It differs from the
XX CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the
XX CC presence of an additional 321 amino acids at the N-terminus (the cDNA
XX CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP
XX CC polypeptides of the invention that are capable of binding to native and
XX CC methylated low density lipoproteins. Also claimed are biologically active
XX CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
XX CC as expression vectors, cells and methods of producing the LBPs.
XX CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529
XX CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of
XX CC determining if an animal is at risk for atherosclerosis, methods for
XX CC evaluating an agent for use in treating atherosclerosis, and methods for
XX CC treating a cell having an abnormality in structure or metabolism of LBP
XX CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
XX CC nucleic acid, and vaccine compositions, are also claimed
XX CC Sequence 538 AA;
Query Match 100.0%; Score 136; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1 EEEEDDEDEDDVSGSEVPESD 26
|||||
RESULT 6
AAB82816
ID AAB82816 standard; peptide; 28 AA.

```


PF	26-NOV-1997;	97WO-US021857.
XX		
PR	27-NOV-1996;	96US-0031930P.
PR	03-JUN-1997;	97US-0048547P.
XX		
PA	(BOST-) BOSTON HEART FOUND INC.	
XX		
XX	Lees AM, Lees RS, Law SW, Arjona AA;	
XX		
XX	WPI; 1998-322455/28.	
DR	N-PSDB; AAV32835.	
XX		
PT	Nucleic acid encoding low density lipoprotein binding proteins and	
PT	related vectors - transformed cells, proteins, and modulators of binding,	
PT	useful for treatment and diagnosis of atherosclerosis and for identifying	
PT	subjects at risk.	
XX		
PS	Claim 1; Fig 2; 47pp; English.	
XX		
CC	This polypeptide comprises novel rabbit low density lipoprotein (LDL)	
CC	binding protein LBP-2 that is capable of binding both native and methyl	
CC	LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA	
CC	(see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and	
CC	human LBPs (see AAV49037-42) are claimed. An abnormality in an aspect of	
CC	LBP metabolism or structure is diagnostic of a risk for atherosclerosis.	
CC	The invention provides: methods for determining if an animal is at risk	
CC	for atherosclerosis (e.g. for prenatal screening); methods for treating	
CC	atherosclerosis (including gene therapy) using e.g. LBP polypeptides to	
CC	bind LDL and thereby prevent formation of atherosclerotic plaque; and	
CC	methods for treating a cell having an abnormality in LBP structure and	
CC	metabolism. Pharmaceutical and vaccine compositions are also provided, as	
CC	well as recombinant vectors and host cells used to produce recombinant	
XX	LBP	
XX		
SQ	Sequence 317 AA;	
	Query Match	76.8%; Score 104.5; DB 2; Length 317;
	Best Local Similarity	74.1%; Pred. No. 3.9e-05;
	Matches 20; Conservative	6; Mismatches 0; Indels 1; Gaps 1;
QY	1 EEEEDDEDEDEDD-VSEGEVPESD 26	
	: : : : : : : : : :	
DB	106 EEEEEDEDEDDDDVVSEGEVPESD 132	
RESULT 10		
AA82798		
ID	AA82798 standard; protein; 317 AA.	
XX		
AC	AA82798;	
AC		
DT	12-NOV-2001 (first entry)	
XX		
DE	Rabbit low density lipoprotein binding protein 2 (LBP-2).	
XX		
KW	Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;	
KW	atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.	
XX		
OS	Oryctolagus cuniculus.	
XX		
XX	Key	Location/Qualifiers
FT	Misc-difference 10	/note= "encoded by TAG"
XX		
PN	WC200164874-A2.	
XX		
PD	07-SEP-2001.	
XX		
XX	28-FEB-2001; 2001WO-US006356.	
PF		
XX	02-MAR-2000; 2000US-00517849.	
PR	14-JUL-2000; 2000US-00616289.	
XX		

CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerray; immunomodulator;
CC antiidiabetic; antiaethmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antihydroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nontropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 2.92778 Seconds
(without alignments)
854.447 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEEDDEDEDDVSEGVSPED 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	64.7	599	S18735	centromere protein
2	83.5	61.4	305	T52287	probable histone d
3	82	60.3	358	S60892	nucleosome assembl
4	82	60.3	606	S70358	centromere protein
5	82	60.3	678	A54514	glutamic acid-rich
6	82	60.3	712	JH0148	nucleolin - rat
7	81	59.6	694	DNCHNL	nucleolin - chicke
8	81	59.6	1161	I59311	nardilysin (EC 3.4
9	80	58.8	707	A35804	nucleolin - human
10	79	58.1	299	A29681	nucleophosmin - Af
11	79	58.1	542	T48488	related to Che-1 p
12	79	58.1	636	T51893	hypothetical prote
13	79	58.1	802	S48529	NAB3 protein - yea
14	79	58.1	1089	S48244	NMD2 protein - yea
15	79	58.1	1105	T18295	Ap-3 adaptor compl
16	79	58.1	2206	G71611	hypothetical prote
17	78	57.4	365	F71437	probable resistanc
18	78	57.4	562	S38149	SIS2 protein - yea
19	78	57.4	792	T42963	hypothetical prote
20	77.5	57.0	1102	H84545	probable ubiquitin
21	77	56.6	896	T24169	hypothetical prote
22	77	56.6	906	T24166	hypothetical prote
23	76.5	56.2	1085	S55352	IFH1 protein - yea
24	76	55.9	677	S15667	transcription fact
25	76	55.9	864	T30441	probable capsid-as
26	76	55.9	1948	S00485	Gene 11-1 protein
27	75.5	55.5	727	S18193	autoantigen NOR-90
28	75.5	55.5	764	S09318	transcription fact
29	75	55.1	279	S30766	ASF1 protein - yea

30	75	55.1	280	2	A61047	ectodermal (ect) -
31	75	55.1	299	2	T02515	cytoskeletal prote
32	75	55.1	1132	2	T43483	translation initia
33	74.5	54.8	727	2	JCS113	ribosomal transcri
34	74.5	54.8	764	2	JCS112	ribosomal transcri
35	74.5	54.8	1133	2	T12529	hypothetical prote
36	74	54.4	121	2	T12467	hypothetical prote
37	74	54.4	147	1	RDBYUC	ubiquinol-cytochro
38	74	54.4	288	2	T22846	hypothetical prote
39	74	54.4	527	2	T18232	conserved hypochet
40	74	54.4	757	1	T50959	probable endopolyp
41	74	54.4	764	2	S64951	hypothetical prote
42	74	54.4	798	2	T33022	hypothetical prote
43	74	54.4	2094	2	S33124	tpr protein - huma
44	73.5	54.0	458	2	D88950	protein R09B5.1 [i
45	73	53.7	215	1	S01947	nonhistone chromos

ALIGNMENTS

RESULT 1

S18735 centromere protein B - human

N;Alternate names: major centromere autoantigen CENP-B

C;Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C;Accession: S18735; A27272

R;Sullivan, K.F.; Glass, C.A.

Chromosoma 100, 360-370, 1991

A;Title: CENP-B is a highly conserved mammalian centromere protein with homology to the

A;Reference number: S18735; MUID:91372020; PMID:1893793

A;Accession: S18735

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-599 <SUL>

A;Cross-references: UNIPROT:P07199; EMBL:X55039; NID:g29860; PIDN:CAA38879.1; PID:g29861

R;Earnshaw, W.C.; Sullivan, K.F.; Machlin, P.S.; Cooke, C.A.; Kaiser, D.A.; Pollard, T.D

J. Cell Biol. 104, 817-829, 1987

A;Title: Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen.

A;Reference number: A27272; MUID:87166180; PMID:2435739

A;Accession: A27272

A;Molecule type: mRNA

A;Residues: 6-582,'M',584-591,'LL',594-599 <EAR>

A;Cross-references: GB:X05299; NID:g29862; PIDN:CAA28918.1; PID:g29863

C;Genetics: GDB:CENPB

A;Gene: GDB:CENPB

A;Cross-references: GDB:118768; OMIM:117140

A;Map position: 20p13-20p13

C;Keywords: DNA binding

Query Match 64.7%; Score 88; DB 2; Length 599;

Best Local Similarity 60.9%; Pred. No. 0.054; Matches 14; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGVSP 23

Db 518 DDEEDDEDDDEDDGDEVP 540

RESULT 2

T52287

probable histone deacetylase (EC 3.5.1.-) [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004

C;Accession: T52287

R;Dangl, M.; Haas, H.; Loidl, P.

submitted to the EMBL Data Library, January 1998

A;Description: Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2.

A;Reference number: Z26016

A;Accession: T52287

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-305 <DAN>
A;Cross-references: UNIPROT:O49209; EMBL:AF044914; PIDN:AAC02539.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: HD2
C;Keywords: hydrolase

Query Match 61.4%; Score 83.5; DB 2; Length 305;
Best Local Similarity 66.7%; Pred. No. 0.081;
Matches 18; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDDVSE-GSEVPSD 26
||:|||||:|||||:|||||:|||||:
Db 156 EEEEDDEDESEDDDDSGKMDVDEDD 182

RESULT 3
S60892
nucleosome assembly protein 1 - soybean
C;Species: Glycine max (soybean)
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S60892
R;Yoon, H.W.; Kim, M.C.; Lee, S.Y.; Hwang, I.; Bahk, J.D.; Hong, J.C.; Ishimi, Y.; Cho, M.I. Gen. Genet. 249, 465-473, 1995
A;Title: Molecular cloning and functional characterization of a cDNA encoding nucleosome
A;Reference number: S60892; MUID:96133687; PMID:8544812
A;Accession: S60892
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <YOO>
A;Cross-references: UNIPROT:Q39892; EMBL:L38856; NID:g1161251; PIDN:AAA88792.1; PID:g116

Query Match 60.3%; Score 82; DB 2; Length 358;
Best Local Similarity 71.4%; Pred. No. 0.13;
Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSE 21
:||||:|||||:|||||:|||||:
Db 307 DEEEDDEDEDEDEDEDE 327

RESULT 4
S70358
centromere protein B - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C;Accession: S70358
R;Bejarano, L.A.; Valdivia, M.M.
Biochim. Biophys. Acta 1307, 21-25, 1996
A;Title: Molecular cloning of an intronless gene for the hamster centromere antigen CENP
A;Reference number: S70358; MUID:96254058; PMID:8652663
A;Accession: S70358
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-606 <BEJ>
A;Cross-references: UNIPROT:P48988; EMBL:U20951; NID:g836955; PIDN:AAB06494.1; PID:g8369
C;Keywords: DNA binding

Query Match 60.3%; Score 82; DB 2; Length 606;
Best Local Similarity 56.5%; Pred. No. 0.22;
Matches 13; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVP 23
|:||||:||||:||||:||||:
Db 525 EDEEEDDEDEDDDDDDGDEVP 547

RESULT 5
A54514
glutamic acid-rich protein precursor - malaria parasite (Plasmodium falciparum)
N;Alternate names: GARP
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 09-Jul-2004

C;Accession: A54514
R;Triglia, T.; Stahl, H.D.; Crewther, P.E.; Silva, A.; Anders, R.F.; Kemp, D.J.
Mol. Biochem. Parasitol. 31, 199-202, 1988
A;Title: Structure of a Plasmodium falciparum gene that encodes a glutamic acid-rich prot
A;Reference number: A54514; MUID:89040048; PMID:2903445
A;Accession: A54514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-678 <TRI>
A;Cross-references: UNIPROT:P13816; GB:J03998; NID:g160298; PIDN:AAA29605.1; PID:g160299
C;Genetics:
A;Introns: 25/3
C;Superfamily: histone H1
C;Keywords: tandem repeat

Query Match 60.3%; Score 82; DB 2; Length 678;
Best Local Similarity 57.7%; Pred. No. 0.25;
Matches 15; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVPESD 26
||:|||||:|||||:|||||:|||||:
Db 640 EEEEDDEDEDEDEDEDEDEDEDEDE 665

RESULT 6
JH0148
nucleolin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C;Accession: JH0148; A24088; I63130
R;Bourbon, H.M.; Amalric, F.
Gene 88, 187-196, 1990
A;Title: Nucleolin gene organization in rodents: highly conserved sequences within three
A;Reference number: JH0148; MUID:90269607; PMID:2347493
A;Accession: JH0148
A;Molecule type: DNA
A;Residues: 1-712 <BOU>
A;Note: the authors translated the initiation codon GTG for residue 1 as Val
R;Lischwe, M.A.; Cook, R.G.; Ahn, Y.S.; Yeoman, L.C.; Busch, H.
Biochemistry 24, 6025-6028, 1985
A;Title: Clustering of glycine and Ng,Ng-dimethylarginine in nucleolar protein C23.
A;Reference number: A24088; MUID:86104094; PMID:4084504
A;Accession: A24088
A;Molecule type: protein
A;Residues: 651-703 <LIS>
R;Bourbon, H.
Gene 68, 73-84, 1988
A;Title: Sequence and structure of the nucleolin promoter in rodents: Characterization of
A;Reference number: I48118; MUID:89121496; PMID:2906027
A;Accession: I63130
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 'MV', 2-44 <RES>
A;Cross-references: GB:M22090; NID:g205793; PIDN:AAA41733.1; PID:g205794
C;Comment: This protein is the major nucleolar-specific protein in eukaryotic exponential

Query Match 60.3%; Score 82; DB 2; Length 712;
Best Local Similarity 57.7%; Pred. No. 0.26;
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVPESD 26
|||||:|||||:|||||:|||||:
Db 241 EEEEDDEDEDEDEDEDEDEDEDD 266

RESULT 7
 DNDCHNL
 nucleolin - chicken
 N/Alternate names: nucleolar protein C23
 C/Species: Gallus gallus (chicken)
 C/Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 09-Jul-2004
 C/Accession: S08414; S10766; A32725; I50397; B30099
 R/Maridor, G.; Nigg, E.A.
 Nucleic Acids Res 18, 1286, 1990
 A/Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar proteins
 A/Reference number: S08414; MUID:90206792; PMID:2320420
 A/Accession: S08414
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-694 <MAR>
 A/Cross-references: UNIPROT:P15771; EMBL:X17199; NID:G63710; PTDN:CAA35060.1; PID:G63710
 R/Maridor, G.; Krek, W.; Nigg, E.A.
 Biochim. Biophys. Acta 1049, 126-133, 1990
 A/Title: Structure and developmental expression of chicken nucleolin and NO38: coordinated
 A/Reference number: S10766; MUID:90304215; PMID:2114180
 A/Accession: S10766
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-694 <MAR2>
 R/Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.
 Cell 60, 791-801, 1990
 A/Title: Identification of major nucleolar proteins as candidate mitotic substrates of cdc2
 A/Reference number: A32725; MUID:90182668; PMID:2178776
 A/Accession: A32725
 A/Molecule type: protein
 A/Residues: 56-62;63-109;197-214 <PET>
 R/Borer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigg, E.A.
 Cell 56, 379-390, 1989
 A/Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.
 A/Reference number: I50397
 A/Accession: I50397
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 407-418, 'R', 420-519, 'T', 521-694 <BOR>
 A/Cross-references: GB:M21791; NID:G212411; PTDN:AAA48983.1; PID:G212412
 C/Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorganization of nucleolin; ribonucleoprotein repeat homology
 C/Superfamily: nucleolin; ribonucleoprotein repeat homology
 C/Keywords: DNA binding; duplication; nucleolus; nucleus; phosphoprotein; RNA binding; RNA binding; RNA binding; RNA binding
 F:1-247/Domain: acidic <AC1>
 F:54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])
 F:254-262/Region: nuclear location signal
 F:267-275/Region: nuclear location signal
 F:282-347/Domain: ribonucleoprotein repeat homology <RRM1>
 F:283-361/Domain: RNA binding #status predicted <RNA1>
 F:372-435/Domain: ribonucleoprotein repeat homology <RRM2>
 F:373-448/Domain: RNA binding #status predicted <RNA2>
 F:462-525/Domain: ribonucleoprotein repeat homology <RRM3>
 F:463-538/Domain: RNA binding #status predicted <RNA3>
 F:554-618/Domain: ribonucleoprotein repeat homology <RRM4>
 F:555-631/Domain: RNA binding #status predicted <RNA4>
 F:632-694/Domain: glycine/arginine-rich <GRR>
 F:56,63,70,77,85/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predicted
 Query Match 59.6%; Score 81; DB 1; Length 694;
 Best Local Similarity 83.3%; Pred. No. 0.32;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEEEDDEDEDEDDVSE 18
 |||||:||||:|||||
 Db 173 EEEEDDEDEDEDDSE 190
 RESULT 8
 I59311
 N/Alternate names: N-arginine dibasic (NRD) convertase
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C/Accession: I59311
 R/Pierotti, A.R.; Prat, A.; Chesneau, V.; Gaudoux, F.; Leseney, A.M.; Foulon, T.; Cohen, Proc. Natl. Acad. Sci. U.S.A. 91, 6078-6082, 1994
 A/Title: N-arginine dibasic convertase, a metalloendopeptidase as a prototype of a class
 A/Reference number: I59311; MUID:94286580; PMID:8016118
 A/Accession: I59311
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA; protein
 A/Residues: 1-1161 <PIB>
 A/Cross-references: UNIPROT:P47245; GB:I27124; NID:G529591; PTDN:AAA21818.1; PID:G529592
 A/Note: the source is designated as Rattus sp. in GenBank entry RATNRDC, release 116.0
 C/Genetics:
 A/Gene: NRDC
 C/Function:
 A/Description: catalyzes the hydrolysis of peptides on the amino side of arginine-arginine
 A/Pathway: protein degradation
 C/Superfamily: insulin-degrading enzyme (IDE)
 C/Keywords: hydrolase; metalloproteinase; protein degradation; zinc
 F:244,248/Binding site: zinc (His) #status predicted
 F:247/Active site: Glu #status predicted
 Query Match 59.6%; Score 81; DB 2; Length 1161;
 Best Local Similarity 46.2%; Pred. No. 0.53;
 Matches 12; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EEEEDDEDEDEDDVSEGEVPEPSD 26
 |||||:||||:|||||:|||||:|||||
 Db 152 EEEEDDEDEDDDDDDSDGARIQDDD 177
 RESULT 9
 A35804
 nucleolin - human
 N/Alternate names: phosphoprotein ppl00; protein B50; protein C23
 C/Species: Homo sapiens (man)
 C/Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text_change 09-Jul-2004
 C/Accession: A35804; S04631; A48138; A55996
 R/Srivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
 J. Biol. Chem. 265, 14922-14931, 1990
 A/Title: Genomic organization and chromosomal localization of the human nucleolin gene.
 A/Reference number: A35804; MUID:90368666; PMID:2394707
 A/Accession: A35804
 A/Molecule type: DNA
 A/Residues: 1-707 <SRI>
 A/Cross-references: UNIPROT:P19338; GB:M60858; GB:J05584; NID:G189305; PTDN:AAA59954.1; I
 R/Srivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
 FEBS Lett. 250, 99-105, 1989
 A/Title: Cloning and sequencing of the human nucleolin cDNA.
 A/Reference number: S04631; MUID:89290043; PMID:2737305
 A/Accession: S04631
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-707 <SR2>
 R/Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.
 Mol. Cell. Biol. 13, 4301-4310, 1993
 A/Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and t
 A/Reference number: A48138; MUID:93309464; PMID:8321232
 A/Accession: A48138
 A/Molecule type: protein
 A/Residues: 458-474 <ISH>
 A/Experimental source: HeLa cell nuclei
 A/Note: sequence extracted from NCBI backbone (NCBI:P134645)
 R/Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.
 Biochemistry 33, 14696-14706, 1994
 A/Title: Major cell surface-located protein substrates of an ecto-protein kinase are hom
 A/Reference number: A55996; MUID:95086063; PMID:7993898
 A/Accession: A55996
 A/Molecule type: protein
 A/Residues: 231-236;349-362;399-403;458-461;655-656, 'X', 658-660 <JOR>
 A/Experimental source: surface-labelled HeLa cells
 C/Genetics:

A;Gene: GDB:NCL
A;Cross-references: GDB:125908; OMIM:164035
A;Map position: 2q12-2qter
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleolus; nucleus; phosphoprotein; RNA binding
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F;487-550/Domain: ribonucleoprotein repeat homology <RRM3>
F;573-634/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 58.8%; Score 80; DB 2; Length 707;
Best Local Similarity 53.8%; Pred. No. 0.41;
Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
:|||||:|||||:|
Db 242 DEDEDDDDDDDDDDDDDEEEEEE 267

RESULT 10
A29681
nucleophosmin - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: A29681
R;Schmidt-Zachmann, M.S.; Huegle-Doerr, B.; Franke, W.W.
EMBO J. 6, 1881-1890, 1987
A;Title: A constitutive nucleolar protein identified as a member of the nucleoplasmin family
A;Reference number: A29681; MUID:88004394; PMID:3308448
A;Accession: A29681
A;Molecule type: mRNA
A;Residues: 1-299 <SCH>
A;Cross-references: UNIPROT:P07222; GB:X05496; NID:g64924; PIDN:CAA29046.1; PID:g64925
C;Superfamily: nucleophosmin
C;Keywords: molecular chaperone; nucleus

Query Match 58.1%; Score 79; DB 2; Length 299;
Best Local Similarity 56.5%; Pred. No. 0.23;
Matches 13; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVP 23
:|||||:|||||:|
Db 167 DEDDDDDDDDDDDEEREETP 189

RESULT 11
T48488
hypothetical protein T28J14.110 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48488
R;Sevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24493
A;Accession: T48488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-542 <BEV>
A;Cross-references: UNIPROT:Q9LYP8; EMBL:AL163652
A;Experimental source: cultivar Columbia; BAC clone T28J14
C;Genetics:
A;Map position: 5
A;Introns: 162/3; 227/1; 251/3; 299/2; 355/3; 377/2; 397/3; 419/3; 435/3; 449/2; 487/3;
A;Note: T28J14.110

Query Match 58.1%; Score 79; DB 2; Length 542;
Best Local Similarity 46.2%; Pred. No. 0.41;
Matches 12; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
:|||||:|||||:|
Db 129 DDDDDDDDDDDDDDDSKSEVEEEE 154

RESULT 12
T51893
related to Che-1 protein [imported] - Neurospora crassa
N;Alternate names: protein B23111.50
C;Species: Neurospora crassa
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T51893
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25858
A;Accession: T51893
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-636 <SCH>
A;Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23111.50
A;Experimental source: BAC clone B23111; strain OR74A
C;Genetics:
A;Gene: NCSP:B23111.50
A;Map position: 6 59/2
A;Introns: 24/1; 59/2

Query Match 58.1%; Score 79; DB 2; Length 636;
Best Local Similarity 50.0%; Pred. No. 0.47;
Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
:|||||:|||||:|
Db 95 DEEEDDEDDDDLEDGESETGSE 120

RESULT 13
S48529
NAB3 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: probable RNA/ssDNA-binding protein HMD1; protein P1945; protein YPL190c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S48529; S65209; S60122
R;Wilson, S.M.; Oberdorfer, A.M.; Datar, K.V.; Swedlow, J.R.; Paddy, M.R.; Swanson, M.S.
submitted to the EMBL Data Library, January 1994
A;Description: Characterization of Nuclear Polyadenylated RNA-Binding Proteins from Sacca
A;Reference number: S48529
A;Accession: S48529
A;Molecule type: DNA
A;Residues: 1-802 <WIL>
A;Cross-references: UNIPROT:P38996; EMBL:U05314; NID:g476219; PID:g476220
R;Rieger, M.; Mueller-Auer, S.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65202
A;Accession: S65209
A;Molecule type: DNA
A;Residues: 1-802 <RIE>
A;Cross-references: EMBL:Z73546; NID:g1370396; PID:g1370397; GSPDB:GN00016; MIPS:YPL190c
A;Experimental source: strain S288C (AB972)
R;Sugimoto, K.; Matsumoto, K.; Kornberg, R.D.; Reed, S.I.; Wittenberg, C.
Mol. Gen. Genet. 248, 712-718, 1995
A;Title: Dosage suppressors of the dominant G1 cyclin mutant CLN3-2: identification of a
A;Reference number: S60122; MUID:96069710; PMID:7476874
A;Accession: S60122
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-340, 1', 342-802 <SUG>
A;Cross-references: GB:D37935; NID:g1235749; PID:g1235750
C;Genetics:
A;Gene: SGD:NAB3; NAB3; HMD1; MIPS:YPL190c
A;Cross-references: SGD:S0006111
A;Map position: 16L
C;Superfamily: yeast NAB3 protein; ribonucleoprotein repeat homology
C;Keywords: nucleus; RNA binding
F;331-396/Domain: ribonucleoprotein repeat homology <RRM1>

Query Match 58.1%; Score 79; DB 1; Length 802;
Best Local Similarity 53.8%; Pred. No. 0.59;

T18295
Ap-3 adaptor complex beta3A chain - mouse
C/Species: Mus musculus (house mouse)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18295
R/Feng, L.; Seymour, A.B.; Jiang, S.Y.; To, A.; Peden, A.A.; Novak, E.K.; Zhen, L.; Rusi
Hum. Mol. Genet. 8, 323-330, 1999
A/Title: The beta3A subunit gene (Ap3b1) of the AP-3 adaptor complex is altered in the m
A/Reference number: Z18864; MUID:99135912; PMID:9931340
A/Accession: T18295
A/Status: preliminary; translated from GB/EMBL/DBJ

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 13.2069 Seconds
(without alignments)
1008.117 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEEDDEDEDDVSEGSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	236	2 Q961M4	Q961M4 homo sapien
2	136	100.0	285	2 Q6P0R3	Q6P0R3 homo sapien
3	136	100.0	295	2 Q6P1S7	Q6P1S7 homo sapien
4	136	100.0	538	2 Q6SPF0	Q6SPF0 homo sapien
5	104.5	76.8	550	2 Q6SPF9	Q6SPF9 oryctolagus
6	97	71.3	239	1 CENB SHEEP	P49451 ovis aries
7	88	64.7	138	2 Q96E14	Q96E14 homo sapien
8	88	64.7	599	1 CENB HUMAN	P07199 homo sapien
9	87	64.0	1161	1 NRDC_MOUSE	Q8bhq1 mus musculus
10	87	64.0	1161	2 Q8R320	Q8R320 mus musculus
11	84	61.8	490	2 Q81E75	Q81E75 plasmodium
12	84	61.8	976	2 Q9DUN0	Q9DUN0 human herpe
13	84	61.8	1036	2 Q9DUM3	Q9DUM3 human herpe
14	84	61.8	2246	2 Q87253	Q87253 dictyosteli
15	83.5	61.4	305	2 Q49209	Q49209 arabidopsis
16	83.5	61.4	306	2 Q9FNJ6	Q9FNJ6 arabidopsis
17	83	61.0	300	2 Q7R5A5	Q7R5A5 plasmodium
18	83	61.0	374	2 Q7P1T7	Q7P1T7 anophelies g
19	83	61.0	374	2 Q70Z19	Q70Z19 nicotiana t
20	83	61.0	457	2 Q6P440	Q6P440 homo sapien
21	83	61.0	492	2 Q6DKH9	Q6DKH9 homo sapien
22	83	61.0	1158	2 Q7Q7D6	Q7Q7D6 anophelies g
23	83	61.0	1544	2 Q6ZNL7	Q6ZNL7 homo sapien
24	83	61.0	2004	1 CHDB_HUMAN	Q9hck8 homo sapien
25	83	61.0	2275	2 Q81BX2	Q81BX2 plasmodium
26	83	61.0	2302	2 Q68DQ0	Q68DQ0 homo sapien
27	82	60.3	235	2 Q7RCY6	Q7RCY6 plasmodium
28	82	60.3	343	2 Q9W3B2	Q9W3B2 drosophila
29	82	60.3	358	2 Q39892	Q39892 glycine max
30	82	60.3	484	2 Q7RH75	Q7RH75 plasmodium
31	82	60.3	606	1 CENB_CRIGR	P48988 cricetus

RESULT 1

ID	Q961M4	PRELIMINARY;	PRT;	236 AA.
AC	Q961M4;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)			
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)			
DE	LOC90378 protein (Hypothetical protein) (Fragment).			
GN	Name=LOC90378;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymph;			
RX	MEDLINE=22388557; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymph;			
RA	Strausberg R.;			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymph;			
RA	Director MGC Project;			
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC007384; AAH07384.2; -			
DR	EMBL; BC080588; AAH080588.1; -			
DR	HSSP; P39769; 1KW4.			
DR	InterPro; IPR001660; SAM.			
DR	InterPro; IPR010993; SAM_homology.			
DR	Pfam; PF00536; SAM_1; 1.			
DR	SMART; SM00454; SAM; 1.			

P13816 plasmodium
P13883 rattus norv
Q9VX49 drosophila
Q8T245 dictyosteli
P27790 mus musculu
Q7T898 mus musculu
Q923C5 mus musculu
Q6khk7 mycoplasma
Q9puk9 gallus galli
Q9yh06 gallus galli
Q6iq75 brachydanio
Q7zuz8 brachydanio
Q9ypa9 human herpe
P15771 gallus galli

ALIGNMENTS

```

DR PROSITE: PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 136; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVPESD 26
   |||||
Db 27 EEEEDDEDEDDVSEGSEVPESD 52
   |||||

RESULT 2
Q6P0R3 PRELIMINARY; PRT; 285 AA.
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madañ A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH5477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;

Query Match 100.0%; Score 136; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVPESD 26
   |||||
Db 76 EEEEDDEDEDDVSEGSEVPESD 101
   |||||

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madañ A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F95F943C7A696F19 CRC64;

Query Match 100.0%; Score 136; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVPESD 26
   |||||
Db 86 EEEEDDEDEDDVSEGSEVPESD 111
   |||||

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madañ A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AAH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;

Query Match 100.0%; Score 136; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSEVPESD 26
   |||||
Db 76 EEEEDDEDEDDVSEGSEVPESD 101
   |||||

```



```

RA  BURKIN D.J., JONES C.A., BURKIN H.R., MCGREW J.A., MCGREW J.A., BROAD T.E.;
RT  "Sheep CENPB and CENPC genes show a high level of sequence similarity
RL  and conserved synteny with their human homologs.";
RL  Cytogenet. Cell Genet. 74:86-89(1996).
CC  -1- FUNCTION: Interacts with centromeric heterochromatin in
CC  chromosomes and binds to a specific subset of alphoid satellite
CC  DNA, called the CENP-B box. May organize arrays of centromere
CC  satellite DNA into a higher order structure which then directs
CC  centromere formation and kinetochore assembly in mammalian
CC  chromosomes (By similarity).
CC  -1- SUBUNIT: Homodimer (By similarity).
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- SIMILARITY: Contains 1 CENPB domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC  or send an email to license@ebi.ac.uk).
CC  -----
DR  EMBL: U35655; AAA79098.1; -.
DR  InterPro: IPR004875; CENP-B.
KW  Centromere; Chromosomal protein; DNA-binding; Nuclear protein.
FT  NON TER 1
SQ  SEQUENCE 239 AA; 26436 MW; 259C6C72E7D9C135 CRC64;

Query Match 71.3%; Score 97; DB 1; Length 239;
Best Local Similarity 78.3%; Pred. No. 0.025;
Matches 18; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 EEEEDDDDEDDVSEGVSEVP 23
Db  158 EEEEDDDDEDDDEEDDEEDV 180

RESULT 7
Q96E14 PRELIMINARY; PRT; 138 AA.
AC  Q96E14;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Hypothetical protein (fragment).
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Placenta;
RA  Strausberg R.;
RL  Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL: BC012297; AAH12297.1; -.
DR  HSSP; P07199; IUFF.
KW  Hypothetical protein.
FT  NON TER 1
SQ  SEQUENCE 138 AA; 15273 MW; 5287985028B9D41D CRC64;

Query Match 64.7%; Score 88; DB 2; Length 138;
Best Local Similarity 60.9%; Pred. No. 0.11;
Matches 14; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY  1 EEEEDDDDEDDVSEGVSEVP 23
Db  57 DDEEDDDDEDDDEEDDEEDV 79

RESULT 8
CENB_HUMAN STANDARD; PRT; 599 AA.
ID  CENB_HUMAN
AC  P07139;

```

DT 01-APR-1988 (Rel. 07, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Major centromere autoantigen B (Centromere protein B) (CENP-B).
 GS Name=CENPB;
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91372020; PubMed=1893793;
 RA Sullivan K.F., Glass C.A.;
 RT "CENP-B is a highly conserved mammalian centromere protein with
 RT homology to the helix-loop-helix family of proteins.";
 RL Chromosoma 100:360-370(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;
 RA Deloukas P., Matthews L.H., Ashurst J.L., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lechaeisaiho M.H., Levensha M.A., Lloyd D.M., Lovell J.D.,
 RA Marsh V.B., Martin S.L., McConachie L.J., McIlroy K., McKerron T.,
 RA Milne S.A., Mistry D., Moore M.C.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.M., Ramey H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 6-599 FROM N.A.
 RX MEDLINE=87166180; PubMed=2435739; DOI=10.1083/jcb.104.4.817;
 RA

RA Earnshaw W.C., Sullivan K.F., Machlin P.S., Cooke C.A., Kaiser D.A.,
 RA Pollard T.D., Rothfield N.F., Cleveland D.W.;
 RT "Molecular cloning of cDNA for CENP-B, the major human centromere
 RT autoantigen.";
 RL J. Cell Biol. 104:817-829(1987).
 RN [5]
 RP SUBUNITS, AND DOMAINS.
 RX MEDLINE=93107144; PubMed=1469042; DOI=10.1083/jcb.119.6.1413;
 RA Yoda K., Kitagawa K., Masumoto H., Muro Y., Okazaki T.;
 RT "A human centromere protein, CENP-B, has a DNA binding domain
 RT containing four potential alpha helices at the NH2 terminus, which is
 RT separable from dimerizing activity.";
 RL J. Cell Biol. 119:1413-1427(1992).
 RN [6]
 RP STRUCTURE BY NMR OF 1-56.
 RX MEDLINE=98119825; PubMed=9451007; DOI=10.1093/emboj/17.3.827;
 RA Iwahara J., Kigawa T., Kitagawa K., Masumoto H., Okazaki T.,
 RA Yokoyama S.;
 RT "A helix-turn-helix structure unit in human centromere protein B
 RT (CENP-B).";
 RL EMBO J. 17:827-837(1998).
 CC -!- FUNCTION: Interacts with centromeric heterochromatin in
 CC chromosomes and binds to a specific subset of aliphoid satellite
 CC DNA, called the CENP-B box. May organize arrays of centromere
 CC satellite DNA into a higher order structure which then directs
 CC centromere formation and kinetochore assembly in mammalian
 CC chromosomes.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Contains 1 CENPB domain.
 CC -!- SIMILARITY: Contains 1 HTH psq-type DNA-binding domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X55039; CAA38879.1; -.
 DR EMBL; AL109804; CAC17547.1; -.
 DR EMBL; BC053847; AAH53847.1; -.
 DR EMBL; X05299; CAA28918.1; -.
 DR FTR; S18735; S18735.
 DR PDB; 1BW6; NMR; A=1-56.
 DR PDB; 1HLV; X-ray; A=1-131.
 DR PDB; 1UFI; X-ray; A/B/C/D=536-599.
 DR Genew; HGNC:1852; CENPB.
 DR MIM; 117140; -.
 DR GO; GO:0000775; C:chromosome, pericentric region; NAS.
 DR GO; GO:0003682; F:chromatin binding; NAS.
 DR GO; GO:0003696; F:satellite DNA binding; NAS.
 DR GO; GO:0000069; P:centromere/kinetochore complex maturation; NAS.
 DR InterPro; IPR004875; CENP-B.
 DR InterPro; IPR006695; CENP-B_N.
 DR InterPro; IPR006600; CENPB.
 DR InterPro; IPR009057; Homeodomain-like.
 DR Pfam; PF04218; CENP-B_N; 1.
 DR Pfam; PF03184; DDE; 1.
 DR SMART; SM00674; CENPB; 1.
 DR PROSITE; PS50960; HTH_PSQ; 1.
 DR 3D-structure; Centromere; Chromosomal protein; DNA-binding;
 KW Nuclear protein.
 FT DOMAIN 1 52 HTH psq-type.
 FT DNA_BIND 28 48 H-T-H motif.
 FT DNA_BIND 97 129 H-T-H motif.
 FT DOMAIN 404 465 Glu-rich (acidic).
 FT DOMAIN 508 538 Asp/Glu-rich (acidic).
 FT CONFLICT 583 583 R -> M (in Ref. 4).
 FT CONFLICT 592 593 VR -> LL (in Ref. 4).
 FT HELIX 10 22
 FT TURN 24 25

RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";

[illegible]

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Latent nuclear antigen (Fragment).
 OS Human herpesvirus 8 (HHV-8) (Kaposi's sarcoma-associated herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=37296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21165304; PubMed=11264383;
 RY DOI=10.1128/JVI.75.8.3948-3959.2001;
 RA Plozet T., Tramier M., Coppey M., Nicolas J.C., Marechal V.;
 RT "Close but distinct regions of human herpesvirus 8 latency-associated
 nuclear antigen 1 are responsible for nuclear targeting and binding to
 human mitotic chromosomes.";
 RL J. Virol. 75:3948-3959(2001).
 DR EMBL; AF305694; AAG27458.1; -.
 DR HSP; P02649; INPN.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR000533; Tropomyosin.
 DR PRINTS; PR00194; TROPOMYOSIN.
 FT NON TR 1036 1036
 SQ SEQUENCE 1036 AA; 119328 MW; 2959EDD2C1C042B8 CRC64;

Query Match 61.8%; Score 84; DB 2; Length 1036;
 Best Local Similarity 61.8%; Pred. No. 2.1;
 Matches 16; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGVESPD 26
 |||:|||||:|:|:|
 Db 402 EEEEDDEDEDEDDVSEGVESPD 427

RESULT 14

ID Q8T253 PRELIMINARY; PRT; 2246 AA.
 AC Q8T253;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger L., Szatranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RL Nature 418:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC115683; AAL92693.2; -.
 DR InterPro; IPR002110; ANK.
 DR SMART; SM00248; ANK; 5.
 KW Hypothetical protein.
 SQ SEQUENCE 2246 AA; 265604 MW; E99C27A2E19FC67A CRC64;

Query Match 61.8%; Score 84; DB 2; Length 2246;
 Best Local Similarity 57.7%; Pred. No. 4.4;
 Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGVESPD 26
 |||:|||||:|:|:|
 Db 330 EEEEDDEDEDEDDVSEGVESPD 355

RESULT 15

O49209
 ID O49209 PRELIMINARY; PRT; 305 AA.
 AC O49209;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Putative histone deacetylase.
 GN Name=HD2; Synonyms=HD2B;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dangl M., Haas H., Loidl P.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20252855; PubMed=10792817;
 RA Wu K., Tian L., Malik K., Brown D., Miki B.;
 RT "Functional analysis of HD2 histone deacetylase homologues in
 Arabidopsis thaliana.";
 RL Plant J. 22:19-27(2000).
 DR EMBL; AF044914; AAC02539.1; -.
 DR EMBL; AF195546; AAG28473.1; -.
 DR PIR; T52287; T52287.
 DR PIR; T52287; T52287.
 SQ SEQUENCE 305 AA; 32219 MW; 58AA6D5764930944 CRC64;
 Query Match 61.4%; Score 83.5; DB 2; Length 305;
 Best Local Similarity 66.7%; Pred. No. 0.7;
 Matches 18; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
 QY 1 EEEEDDEDEDDVSE-GSEVPESD 26
 |||:|||||:|:|:|
 Db 156 EEEEDDEDEDDVSE-GSEVPESD 182
 Search completed: September 20, 2005, 12:41:44
 Job time : 14.2069 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 14.3525 Seconds
(without alignments)
733.538 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEEDDEDEDDVSRGSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	136	100.0	26	9	US-09-962-055-20 Sequence 20, Appl
2	136	100.0	26	9	US-09-976-740-20 Sequence 20, Appl
3	136	100.0	26	13	US-10-023-523-20 Sequence 20, Appl
4	136	100.0	26	13	US-10-023-523-20 Sequence 20, Appl
5	136	100.0	26	15	US-10-616-187-20 Sequence 20, Appl
6	136	100.0	26	15	US-10-671-242-20 Sequence 20, Appl
7	136	100.0	217	9	US-09-962-055-7 Sequence 7, Appli
8	136	100.0	217	9	US-09-976-740-7 Sequence 7, Appli
9	136	100.0	217	13	US-10-023-523-7 Sequence 7, Appli
10	136	100.0	217	13	US-10-023-523-7 Sequence 7, Appli
11	136	100.0	217	15	US-10-616-187-7 Sequence 7, Appli

12	136	100.0	217	15	US-10-671-242-7	Sequence 7, Appli
13	136	100.0	241	9	US-09-925-298-665	Sequence 665, App
14	136	100.0	241	14	US-10-102-806-665	Sequence 665, App
15	136	100.0	538	9	US-09-976-740-43	Sequence 43, Appl
16	136	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
17	136	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
18	136	100.0	538	15	US-10-616-187-43	Sequence 43, Appl
19	136	100.0	538	15	US-10-671-242-43	Sequence 43, Appl
20	104.5	76.8	28	9	US-09-962-055-26	Sequence 26, Appl
21	104.5	76.8	28	9	US-09-976-740-26	Sequence 26, Appl
22	104.5	76.8	28	13	US-10-023-523-26	Sequence 26, Appl
23	104.5	76.8	28	13	US-10-023-523-26	Sequence 26, Appl
24	104.5	76.8	28	15	US-10-616-187-26	Sequence 26, Appl
25	104.5	76.8	28	15	US-10-671-242-26	Sequence 26, Appl
26	104.5	76.8	232	9	US-09-962-055-3	Sequence 3, Appli
27	104.5	76.8	232	9	US-09-976-740-3	Sequence 3, Appli
28	104.5	76.8	232	13	US-10-023-523-3	Sequence 3, Appli
29	104.5	76.8	232	13	US-10-023-523-3	Sequence 3, Appli
30	104.5	76.8	232	15	US-10-616-187-3	Sequence 3, Appli
31	104.5	76.8	232	15	US-10-671-242-3	Sequence 3, Appli
32	104.5	76.8	252	9	US-09-962-055-4	Sequence 4, Appli
33	104.5	76.8	252	9	US-09-976-740-4	Sequence 4, Appli
34	104.5	76.8	252	13	US-10-023-523-4	Sequence 4, Appli
35	104.5	76.8	252	13	US-10-023-523-4	Sequence 4, Appli
36	104.5	76.8	252	15	US-10-616-187-4	Sequence 4, Appli
37	104.5	76.8	317	9	US-09-962-055-2	Sequence 2, Appli
38	104.5	76.8	317	9	US-09-976-740-2	Sequence 2, Appli
39	104.5	76.8	317	13	US-10-023-523-2	Sequence 2, Appli
40	104.5	76.8	317	13	US-10-023-523-2	Sequence 2, Appli
41	104.5	76.8	317	15	US-10-616-187-2	Sequence 2, Appli
42	104.5	76.8	317	15	US-10-671-242-2	Sequence 2, Appli
43	104.5	76.8	550	9	US-09-976-740-47	Sequence 47, Appl
44	104.5	76.8	550	13	US-10-023-523-47	Sequence 47, Appl
45	104.5	76.8	550	13	US-10-023-523-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-09-962-055-20

; Sequence 20, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA: US/09/962,055

; APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 136; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 2
US-09-976-740-20
; Sequence 20, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-20

Query Match      100.0%; Score 136; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 3
US-10-023-529-20
; Sequence 20, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 136; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 4
US-10-023-523-20
; Sequence 20, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 136; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 1 EEEEDDEDEDDVSGSEVPESD 26
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 136; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 4
US-10-023-523-20
; Sequence 20, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 136; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 1 EEEEDDEDEDDVSGSEVPESD 26
```


RESULT 5
US-10-616-187-20
; Sequence 20, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-20

Query Match 100.0%; Score 136; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 6
US-10-671-242-20
; Sequence 20, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-20

Query Match 100.0%; Score 136; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 7
US-10-671-242-20
; Sequence 7, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match 100.0%; Score 136; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 8
US-09-976-740-7

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-20

Query Match 100.0%; Score 136; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 7
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match 100.0%; Score 136; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 8
US-09-976-740-7

```
; Sequence 7, Application US/09976740
; Publication NO. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-7

Query Match      100.0%; Score 136; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
DB 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 9
US-10-023-529-7
; Sequence 7, Application US/10023529
; Publication NO. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 136; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
DB 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 10
US-10-023-523-7
; Sequence 7, Application US/10023523
; Publication NO. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-7

Query Match      100.0%; Score 136; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
DB 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 11
US-10-616-187-7
; Sequence 7, Application US/10616187
; Publication NO. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
```

```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-7

Query Match      100.0%; Score 136; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 8 EEEEDDEDEDEDDVSGSEVPESD 33

RESULT 12
US-10-671-242-7
; Sequence 7, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 136; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 8 EEEEDDEDEDEDDVSGSEVPESD 33

RESULT 13
US-09-925-298-665
; Sequence 665, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 136; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 8 EEEEDDEDEDEDDVSGSEVPESD 33

RESULT 14
US-10-102-806-665
; Sequence 665, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 665
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-665

Query Match      100.0%; Score 136; DB 9; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 32 EEEEDDEDEDEDDVSGSEVPESD 57

RESULT 15
US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
```

; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-43

Query Match 100.0%; Score 136; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 4,1e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
|||||
Db 329 EEEEDDEDEDDVSGSEVPESD 354
|||||

Search completed: September 20, 2005, 12:53:06
Job time : 14.3525 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 6.54345 Seconds
(without alignments)
650.171 Million cell updates/sec

Title: US-10-671-242-21
Perfect score: 54
Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	11	4	AAB82811 Human low
2	54	100.0	12	4	AAB82817 Rabbit lo
3	54	100.0	26	4	AAB82810 Human low
4	54	100.0	28	4	AAB82816 Rabbit lo
5	54	100.0	217	2	AAB82803 Human low
6	54	100.0	217	4	AAB82803 Human low
7	54	100.0	232	4	AAB82799 Rabbit lo
8	54	100.0	241	3	AAB858957 Breast an
9	54	100.0	252	4	AAB82800 Rabbit lo
10	54	100.0	317	2	AAB82803 Rabbit lo
11	54	100.0	317	4	AAB82798 Rabbit lo
12	54	100.0	538	4	AAB82806 Human low
13	54	100.0	550	4	AAB82807 Rabbit lo
14	38	70.4	759	7	ABM85225 Human pro
15	38	70.4	808	8	ADN24118 Bacterial
16	38	70.4	906	2	RAY27126 Amino aci
17	38	70.4	906	4	ABM79220 Human pro
18	38	70.4	906	6	ABU03746 Human exp
19	38	70.4	906	6	ABU03752 Human exp
20	38	70.4	906	6	ABU03755 Human exp
21	38	70.4	906	6	ABU03750 Human exp
22	38	70.4	906	6	ABU03754 Human exp
23	38	70.4	906	6	AAB38431 Human Cbl
24	38	70.4	906	7	ADB80950 RING-SH c
25	38	70.4	906	8	ADJ66658 c-CBL pro

ALIGNMENTS

RESULT 1

AAB82811	26	38	70.4	906	8	ADP46907	Adp46907 Human c-C
ID AAB82811 standard; peptide; 11 AA.	27	38	70.4	906	8	ADP46913	Adp46913 Human c-C
XX AAB82811;	28	38	70.4	906	8	ADP46667	Adp46667 Human c-C
DT 12-NOV-2001 (first entry)	29	38	70.4	906	8	ADP46909	Adp46909 Human c-C
XX Human low density lipoprotein binding protein 2 (LBP-2) peptide.	30	38	70.4	906	8	ADP46905	Adp46905 Human c-C
XX Low density lipoprotein binding protein 2; LBP-2; LDL; human; atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.	31	38	70.4	906	8	ADP46915	Adp46915 Human c-C
KW Homo sapiens.	32	38	70.4	906	8	ADP46911	Adp46911 Human c-C
XX WO200164874-A2.	33	38	70.4	953	4	AAW80204	AAW80204 Human pro
PD 07-SEP-2001.	34	38	70.4	953	6	ABU03751	ABU03751 Human exp
XX 28-FEB-2001; 2001WO-US006356.	35	37	68.5	158	4	AAU65679	AAU65679 Proptonib
XX 02-MAR-2000; 2000US-00517849.	36	37	68.5	158	4	AAU49300	AAU49300 Proptonib
XX 14-JUL-2000; 2000US-00616289.	37	37	68.5	158	6	ABM45819	ABM45819 Proptonib
(BOST-) BOSTON HEART FOUND INC.	38	37	68.5	158	6	ABM62198	ABM62198 Proptonib
XX Lees AM, Lees RS, Law SW, Arjona AA;	39	37	68.5	178	2	AAW33626	AAW33626 Bimeria t
XX WPI; 2001-565505/63.	40	37	68.5	178	7	ADF77369	ADF77369 28-37Kd p
XX New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.	41	37	68.5	181	2	AAW33625	AAW33625 Bimeria t
XX Claim 14(a); Page; 143pp; English.	42	37	68.5	181	7	ADF77368	ADF77368 28-37Kd p
XX The present sequence is that of a peptide comprising amino acid residues 344-354 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs.	43	37	68.5	193	3	AAW04795	AAW04795 Arabidops
XX Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine	44	37	68.5	263	1	AAW93705	AAW93705 Sequence
XX	45	37	68.5	263	2	AAW31584	AAW31584 Bimeria t

CC compositions. Note: the present sequence is not shown in the
 CC specification but is derived from the human LBP-2 sequence given in
 CC figure 7A (see AAB82806)

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0039; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0;

Qy 1 VSEGSSEVPESD 11
 |||||
 Db 1 VSEGSSEVPESD 11

RESULT 2
 AAB82817
 ID AAB82817 standard; peptide; 12 AA.
 XX
 AC AAB82817;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2) peptide.
 XX
 DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 KW
 XX Oryctolagus cuniculus.
 OS
 XX WO200164874-A2.
 PN
 XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.
 XX
 XX 02-MAR-2000; 2000US-00517849.
 PR
 PR 14-JUL-2000; 2000US-00616289.
 XX
 XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 2001-565505/63.
 DR
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 PT
 XX Claim 14(c); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues
 CC 354-365 of novel rabbit low density lipoprotein binding protein 2 (LBP-2,
 CC see AAB82807). Rabbit LBP-2 is an example of claimed LBP polypeptides of
 CC the invention that are capable of binding to native and methylated low
 CC density lipoproteins (LDLs). Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs.
 CC Polypeptides having the present amino acid sequence are among those
 CC claimed. Methods of determining if an animal is at risk for
 CC atherosclerosis, and methods for evaluating an agent for use in treating
 CC atherosclerosis, and methods for treating a cell having an abnormality in
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
 CC compositions. Note: the present sequence is not shown in the
 CC specification but is derived from the rabbit LBP-2 sequence given in
 CC figure 2A (see AAB82807)

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 54; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.0043;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSSEVPESD 11
 |||||
 Db 2 VSEGSSEVPESD 12

RESULT 3
 AAB82810
 ID AAB82810 standard; peptide; 26 AA.
 XX
 AC AAB82810;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.
 XX
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 KW
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.
 XX
 XX 02-MAR-2000; 2000US-00517849.
 PR
 PR 14-JUL-2000; 2000US-00616289.
 XX
 XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 PT
 XX Claim 14(a); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues
 CC 329-354 of novel human low density lipoprotein binding protein 2 (LBP-2,
 CC see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of
 CC the invention that are capable of binding to native and methylated low
 CC density lipoproteins (LDLs). Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs.
 CC Polypeptides having the present amino acid sequence are among those
 CC claimed. Methods of determining if an animal is at risk for
 CC atherosclerosis, methods for evaluating an agent for use in treating
 CC atherosclerosis, and methods for treating a cell having an abnormality in
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
 CC compositions. Note: the present sequence is not shown in the
 CC specification but is derived from the human LBP-2 sequence given in
 CC figure 7A (see AAB82806)

XX SQ Sequence 26 AA;

Query Match 100.0%; Score 54; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0;

Qy 1 VSEGSSEVPESD 11
 |||||
 Db 16 VSEGSSEVPESD 26

RESULT 4
 AAB82816
 ID AAB82816 standard; peptide; 28 AA.
 XX
 AC AAB82816;

PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 PT neurological diseases.
 PS Claim 11; Page 1112; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 CC antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases
 XX
 SQ Sequence 241 AA;

Query Match 100.0%; Score 54; DB 3; Length 241;
 Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;
 Matches 11; Conservative 0; Indels 0; Gaps 0;

Qy 1 VSEGSVEPESD 11
 | | | | | | | | | |
 Db 47 VSEGSVEPESD 57

RESULT 9
 AAB82800
 ID AAB82800 standard; protein; 252 AA.

AC AAB82800;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 3 (LBP-3).
 XX
 KW Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Oryctolagus cuniculus.

PN WQ200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SM, Arjona AA;

XX WPI; 2001-565505/63.

DR N-PSDB; AAB26490.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(d); Fig 4; 143pp; English.

XX

CC The present sequence is that of a partial sequence of novel rabbit low
 CC density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is
 CC deduced from an isolated cDNA clone (see AAB26490). A full-length
 CC sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an
 CC example of claimed polypeptides of the invention, termed LBPs, that are
 CC capable of binding to native and methylated low density lipoproteins.
 CC Also claimed are biologically active fragments and analogues of LBPs,
 CC polynucleotides encoding LBPs, as well as expression vectors, cells and
 CC methods of producing the LBPs. Methods of determining if an animal is at
 CC risk for atherosclerosis, methods for evaluating an agent for use in
 CC treating atherosclerosis, and methods for treating a cell having an
 CC abnormality in structure or metabolism of LBP are also claimed, as are
 CC pharmaceutical compositions comprising an LBP polypeptide or nucleic
 CC acid, and vaccine compositions
 XX
 SQ Sequence 252 AA;

Query Match 100.0%; Score 54; DB 4; Length 252;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSVEPESD 11
 | | | | | | | | | |
 Db 57 VSEGSVEPESD 67

RESULT 10

AAB49038

ID AAB49038 standard; protein; 317 AA.

XX AAB49038;

XX 09-NOV-1998 (first entry)

DE Rabbit low density lipoprotein binding protein LBP-2.

KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
 KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.

XX Oryctolagus cuniculus.

XX Key Location/Qualifiers

FT Misc-difference 10 /note= "encoded by TAG"

FT Protein 66..317 /note= "Claim 1"

FT Protein 86..317 /note= "Claim 1"

FT Peptide 105..132 /note= "Claim 2"

FT Peptide 105..120 /note= "Claim 2"

FT Peptide 121..132 /note= "Claim 2"

FT Peptide 211..220 /note= "Claim 2"

XX WQ9823282-A1.

XX 04-JUN-1998.

XX 26-NOV-1997; 97WO-US021857.

XX 27-NOV-1996; 96US-0031930P.

PR 03-JUN-1997; 97US-0048547P.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SM, Arjona AA;

XX WPI; 1998-322455/28.

DR N-PSDB; AAB32835.

XX

PT Nucleic acid encoding low density lipoprotein binding proteins and
 PT related vectors - transformed cells, proteins, and modulators of binding,
 PT useful for treatment and diagnosis of atherosclerosis and for identifying
 PT subjects at risk.

XX Claim 1; Fig 2; 47pp; English.

XX This polypeptide comprises novel rabbit low density lipoprotein (LDL)
 CC binding protein LBP-2 that is capable of binding both native and methyl
 CC LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA
 CC (see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and
 CC human LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of
 CC LBP metabolism or structure is diagnostic of a risk for atherosclerosis.
 CC The invention provides: methods for determining if an animal is at risk
 CC for atherosclerosis (e.g. for prenatal screening); methods for treating
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
 CC methods for treating a cell having an abnormality in LBP structure or
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
 CC well as recombinant vectors and host cells used to produce recombinant
 CC LBP

XX Sequence 317 AA;

Query Match 100.0%; Score 54; DB 2; Length 317;
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;
 Matches 11; Conservative 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
 DB 122 VSEGSSEVPESD 132

RESULT 11

AAB82798

ID AAB82798 standard; protein; 317 AA.

XX AAB82798;

DT 12-NOV-2001 (first entry)

DE Rabbit low density lipoprotein binding protein 2 (LBP-2).

XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;

XX atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

OS Oryctolagus cuniculus.

PH Key Location/Qualifiers

FT Misc-difference 10

FT /note= "encoded by TAG"

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAH26488.

XX New isolated low density lipoprotein binding polypeptide for treating,

XX diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(b); Fig 2b; 143pp; English.

CC The present sequence is that of a partial sequence of novel rabbit low
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
 CC deduced from an isolated cDNA clone (see AAH26488). Full-length rabbit
 CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed
 CC polypeptides of the invention, termed LBPs, that are capable of binding
 CC to native and methylated low density lipoproteins. Also claimed are
 CC biologically active fragments and analogues of LBPs, polynucleotides
 CC encoding LBPs, as well as expression vectors, cells and methods of
 CC producing the LBPs. Methods of determining if an animal is at risk for
 CC atherosclerosis, methods for evaluating an agent for use in treating
 CC atherosclerosis, and methods for treating a cell having an abnormality in
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
 CC compositions

XX Sequence 317 AA;

Query Match 100.0%; Score 54; DB 4; Length 317;

Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;

Matches 11; Conservative 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11

DB 122 VSEGSSEVPESD 132

RESULT 12

AAB82806

ID AAB82806 standard; protein; 538 AA.

XX AAB82806;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2).

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;

XX atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

OS Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAH26499.

XX New isolated low density lipoprotein binding polypeptide for treating,

XX diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein
 CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the
 CC coding region of isolated genomic DNA (see AAH26499). It differs from the
 CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the
 CC presence of an additional 321 amino acids at the N-terminus (the cDNA
 CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP
 CC polypeptides of the invention that are capable of binding to native and
 CC methylated low density lipoproteins. Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs.

XX Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529

CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of
 CC determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX
 SQ Sequence 538 AA;

Query Match 100.0%; Score 54; DB 4; Length 538;
 Best Local Similarity 100.0%; Pred. No. 0.27; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVEPESD 11
 |||||
 DB 344 VSEGSVEPESD 354

RESULT 13
 AAB82807
 ID AAB82807 standard; protein; 550 AA.

XX AAB82807;

XX 12-NOV-2001 (first entry)

XX Rabbit low density lipoprotein binding protein 2 (LBP-2).

XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Oryctolagus cuniculus.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-555505/63.

XX N-PSDB; AAH26500.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(k); Fig 2a; 143pp; English.

XX The present sequence is that of a full-length sequence of novel rabbit
 CC low density lipoprotein binding protein 2 (LBP-2). The amino acid
 CC sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit
 CC LBP-2 is an example of claimed polypeptides of the invention, termed
 CC LBP, that are capable of binding to native and methylated low density
 CC lipoproteins. Also claimed are biologically active fragments and
 CC analogues of LBPs, polynucleotides encoding LBPs, as well as expression
 CC vectors, cells and methods of producing the LBPs. Polypeptides having
 CC amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-
 CC 18) of the present sequence are claimed. Methods of determining if an
 CC animal is at risk for atherosclerosis, methods for evaluating an agent
 CC for use in treating atherosclerosis, and methods for treating a cell
 CC having an abnormality in structure or metabolism of LBP are also claimed,
 CC as are pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions

XX Sequence 550 AA;

Query Match 100.0%; Score 54; DB 4; Length 550;

Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VSEGSVEPESD 11
 |||||
 DB 355 VSEGSVEPESD 365

RESULT 14

ABM85225

ID ABM85225 standard; protein; 759 AA.

XX AC ABM85225;

XX 18-NOV-2004 (first entry)

XX Human protein sequence hCP34885.

XX Cytostatic; carcinoma; lymphoma; cancer; human.

XX OS Homo sapiens.

XX PN WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.

XX Claim 5; SEQ ID NO 90; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CAP. Note:
 CC This patent is an equivalent to basic patent US2002182586A1, for which no
 CC sequence data was published

XX Sequence 759 AA;

Query Match 70.4%; Score 38; DB 7; Length 759;

Best Local Similarity 77.8%; Pred. No. 3.7e+02;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVEPVE 9

DB 666 VTEGSQVPE 674

RESULT 15

ADN24118

ID ADN24118 standard; protein; 808 AA.

XX AC ADN24118;

XX

Search completed: September 20, 2005, 12:34:45
Job time : 7.54345 secs

DT 02-DEC-2004 (first entry)
DE Bacterial polypeptide #6771.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 6771; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 808 AA;

Query Match 70.4%; Score 38; DB 8; Length 808;
Best Local Similarity 77.8%; Pred. No. 4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 EGSEVPESD 11
||| |:
Db 338 EGSETPKSD 346

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 1.66952 Seconds
(without alignments)
491.841 Million cell updates/sec

Title: US-10-671-242-21
Perfect score: 54
Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	11	3	US-08-979-608A-21
2	54	100.0	11	4	US-09-517-849-21
3	54	100.0	11	4	US-09-616-289-21
4	54	100.0	12	3	US-08-979-608A-27
5	54	100.0	12	4	US-09-517-849-27
6	54	100.0	12	4	US-09-616-289-27
7	54	100.0	26	3	US-08-979-608A-20
8	54	100.0	26	4	US-09-517-849-20
9	54	100.0	26	4	US-09-616-289-20
10	54	100.0	28	3	US-08-979-608A-26
11	54	100.0	28	4	US-09-517-849-26
12	54	100.0	28	4	US-09-616-289-26
13	54	100.0	217	3	US-08-979-608A-7
14	54	100.0	217	4	US-09-517-849-7
15	54	100.0	217	4	US-09-616-289-7
16	54	100.0	232	3	US-08-979-608A-3
17	54	100.0	232	4	US-09-517-849-3
18	54	100.0	232	4	US-09-616-289-3
19	54	100.0	252	3	US-08-979-608A-4
20	54	100.0	252	4	US-09-517-849-4
21	54	100.0	252	4	US-09-616-289-4
22	54	100.0	317	3	US-08-979-608A-2
23	54	100.0	317	4	US-09-517-849-2
24	54	100.0	317	4	US-09-616-289-2
25	54	100.0	538	4	US-09-616-289-43
26	54	100.0	550	4	US-09-616-289-47
27	36	66.7	277	4	US-09-949-016-6068

28	36	66.7	363	4	US-09-107-532A-5091	Sequence 5091, Ap
29	36	66.7	746	4	US-09-248-796A-20280	Sequence 20280, A
30	35.5	65.7	755	4	US-09-134-000C-6691	Sequence 6691, Ap
31	35	64.8	263	4	US-09-134-000C-5721	Sequence 5721, Ap
32	35	64.8	287	4	US-09-489-039A-8878	Sequence 8878, Ap
33	35	64.8	415	4	US-09-252-991A-23602	Sequence 23602, A
34	35	64.8	456	4	US-09-902-540-11227	Sequence 11227, A
35	34	63.0	97	4	US-09-621-976-6690	Sequence 6690, Ap
36	34	63.0	121	3	US-09-134-001C-3834	Sequence 3834, Ap
37	34	63.0	173	4	US-08-956-171B-5218	Sequence 5218, Ap
38	34	63.0	173	4	US-08-781-986A-5218	Sequence 5218, Ap
39	34	63.0	373	4	US-09-710-279-372	Sequence 372, Appl
40	34	63.0	591	4	US-09-602-787A-24	Sequence 24, Appl
41	34	63.0	712	4	US-09-708-426-9	Sequence 9, Appl
42	34	63.0	857	2	US-08-659-251-2	Sequence 2, Appl
43	34	63.0	857	3	US-09-256-490-2	Sequence 2, Appl
44	34	63.0	857	5	PCT-US96-11445-2	Sequence 2, Appl
45	34	63.0	1247	1	US-08-472-934-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-08-979-608A-21
; Sequence 21, Application US/08979608A
; Patent No. 6355451

GENERAL INFORMATION:

APPLICANT: Lees, Ann M. S.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,608A

FILING DATE: 26-Nov-6355451-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,547

FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930

FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 21:

US-08-979-608A-21

Query Match 100.0%; Score 54; DB 3; Length 11;

```
Best Local Similarity 100.0%; Pred. No. 0.00097; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 2
US-09-517-849-21
; Sequence 21, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-517-849-21

Query Match 100.0%; Score 54; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 3
US-09-616-289-21
; Sequence 21, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

Best Local Similarity 100.0%; Score 54; DB 4; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-09-616-289-21

Query Match 100.0%; Score 54; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 4
US-08-979-608A-27
; Sequence 27, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-08-979-608A-27

Query Match      100.0%; Score 54; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSEVPESD 11
Db      1 VSEGSEVPESD 12

RESULT 5
US-09-517-849-27
; Sequence 27, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;                   AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; TYPE: PRT
; LENGTH: 12
; ORGANISM: Oryctolagus cuniculus
;
US-09-616-289-27

Query Match      100.0%; Score 54; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSEVPESD 11
Db      2 VSEGSEVPESD 12

RESULT 7
US-08-979-608A-20
; Sequence 20, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN
;                   DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-517-849-27

Query Match      100.0%; Score 54; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSEVPESD 11
Db      2 VSEGSEVPESD 12

RESULT 6
```

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-979-608A-20

Query Match      100.0%; Score 54; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VSEGSVPESD 11
Db      16 VSEGSVPESD 26

RESULT 8
US-09-517-849-20
; Sequence 20, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20

Query Match      100.0%; Score 54; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
```

```
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VSEGSVPESD 11
Db      16 VSEGSVPESD 26

RESULT 9
US-09-616-289-20
; Sequence 20, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 26
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-616-289-20

Query Match      100.0%; Score 54; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VSEGSVPESD 11
Db      16 VSEGSVPESD 26

RESULT 10
US-08-979-608A-26
; Sequence 26, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20

Query Match      100.0%; Score 54; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
```


/ APPLICATION NUMBER: US/08/979,608A
/ FILING DATE: 26-No. 6355451-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/048,547
/ FILING DATE: 03-JUN-1997
/ APPLICATION NUMBER: US 60/031,930
/ FILING DATE: 27-NOV-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 28 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-979-608A-26

Query Match 100.0%; Score 54; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
DB 18 VSEGSSEVPESD 28

RESULT 11
US-09-517-849-26
/ Sequence 26, Application US/09517849
/ Patent No. 6605588
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ Lees, Robert S.
/ Law, Simon W.
/ Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
/ TREATING ATHEROSCLEROSIS
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/517,849
/ FILING DATE: 02-Mar-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/979,608
/ FILING DATE: 26-NOV-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-003001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 28 amino acids

/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-517-849-26

Query Match 100.0%; Score 54; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
DB 18 VSEGSSEVPESD 28

RESULT 12
US-09-616-289-26
/ Sequence 26, Application US/09616289
/ Patent No. 6632923
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ APPLICANT: Lees, Robert S.
/ APPLICANT: Law, Simon W.
/ APPLICANT: Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
/ PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
/ ATHEROSCLEROSIS
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/09/616,289
/ CURRENT FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
/ PRIOR FILING DATE: 1996-11-27
/ PRIOR APPLICATION NUMBER: US 60/048,547
/ PRIOR FILING DATE: 1997-06-03
/ NUMBER OF SEQ ID NOS: 53
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 28
/ TYPE: PRT
/ ORGANISM: Oryctolagus cuniculus
US-09-616-289-26

Query Match 100.0%; Score 54; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
DB 18 VSEGSSEVPESD 28

RESULT 13
US-08-979-608A-7
/ Sequence 7, Application US/08979608A
/ Patent No. 6355451
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ Lees, Robert S.
/ Law, Simon W.
/ Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
/ TREATING ATHEROSCLEROSIS
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA

```
;
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match 100.0%; Score 54; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSSEVPESD 11
DB 23 VSEGSSEVPESD 33

RESULT 14
US-09-517-849-7
; Sequence 7, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-MAR-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
```

```
;
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match 100.0%; Score 54; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSSEVPESD 11
DB 23 VSEGSSEVPESD 33

RESULT 15
US-09-616-289-7
; Sequence 7, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-7

Query Match 100.0%; Score 54; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSSEVPESD 11
DB 23 VSEGSSEVPESD 33
```

Search completed: September 20, 2005, 12:45:30
Job time : 2.66952 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 1.23868 Seconds
(without alignments)
854.447 Million cell updates/sec

Title: US-10-671-242-21

Perfect score: 54

Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	75.9	244	2 F69066	conserved hypother
2	41	75.9	587	2 E84808	hypothetical prote
3	38	70.4	139	2 AF0221	conserved hypother
4	38	70.4	808	2 T16564	hypothetical prote
5	38	70.4	906	2 A43817	transforming prote
6	38	70.4	1661	2 A83695	hypothetical prote
7	37	68.5	258	2 D95284	probable [imported
8	37	68.5	427	2 S57776	cysteine proteinas
9	37	68.5	447	2 T46146	hypothetical prote
10	37	68.5	831	2 A70363	mannose-1-phosphat
11	37	68.5	2218	2 B84683	hypothetical prote
12	36	66.7	75	2 E87604	conserved hypother
13	36	66.7	204	2 T33362	hypothetical prote
14	36	66.7	496	2 T52112	deoxyribodipyrimid
15	36	66.7	723	2 S51788	malate synthase (E
16	36	66.7	1877	2 T21861	hypothetical prote
17	35	64.8	216	2 H69443	conserved hypother
18	35	64.8	316	2 E89973	hypothetical prote
19	35	64.8	1173	2 T51892	hypothetical prote
20	35	64.8	1883	2 G82875	hypothetical prote
21	35	64.8	2301	2 T02323	nodulin-like prote
22	35	64.8	6420	2 T30283	polyketide synthas
23	34.5	63.9	341	2 S43586	F26F3.4 protein (c
24	34	63.0	164	2 E89955	hypothetical prote
25	34	63.0	183	2 C97832	alpha-(1,3)-fucosyl
26	34	63.0	281	2 JC4295	heat-shock protein
27	34	63.0	352	2 A56123	scraptoirisin D (E
28	34	63.0	513	2 A96265	hypothetical prote
29	34	63.0	513	2 AH3019	sigma 54 dependent

30	34	63.0	551	1 VGN2P2	cell fusion glycop
31	34	63.0	551	1 VGN2PG	cell fusion glycop
32	34	63.0	552	2 T44954	flagella-related p
33	34	63.0	590	2 T42202	probable acyl-CoA
34	34	63.0	590	2 T42206	probable acyl-CoA
35	34	63.0	613	2 T24662	hypothetical prote
36	34	63.0	638	2 T24661	hypothetical prote
37	34	63.0	639	2 B95945	probable glycosyl
38	34	63.0	728	2 AF3299	malate synthase (E
39	34	63.0	819	2 C84615	hypothetical prote
40	34	63.0	877	2 T03098	p97 protein - Toxo
41	34	63.0	896	2 B43817	transforming prote
42	34	63.0	1022	2 T24663	hypothetical prote
43	34	63.0	1039	2 T43678	tetratricopeptide
44	34	63.0	1039	2 T38447	tetratricopeptide
45	-34	63.0	1129	2 T42732	A-kinase anchoring

ALIGNMENTS

RESULT 1

F69066

Conserved hypothetical protein MTH1498 - Methanobacterium thermoautotrophicum (strain Del

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: F69066

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.,
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: F69066

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-244 <MTH>

A;Cross-references: UNIPROT:O27542; GB:AE000910; GB:AE000666; NID:g2622610; PIDN:AB88597;

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1498

A;Start codon: GTG

C;Superfamily: conserved hypothetical protein MTH593

Query Match 75.9%; Score 41; DB 2; Length 244;
Best Local Similarity 54.5%; Pred. No. 3.6;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11

Db 231 ITEGEIPEED 241

RESULT 2

E84808

hypothetical protein At2g38720 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: E84808

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: E84808

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-587 <STO>

A;Cross-references: UNIPROT:Q9ZVJ3; GB:AE002093; NID:g3786000; PIDN:AAC67346.1; GSPDB:GN

C;Genetics:

A;Gene: At2g38720

A;Map position: 2

```
Query Match      75.9%; Score 41; DB 2; Length 587;
Best Local Similarity 81.8%; Pred. No. 9.5;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSEGSVEPESD 11
   |||||
Db 137 VSSGSEVDES 147

RESULT 3
AF0221
conserved hypothetical protein YP01817 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF0221
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0221
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139 <CUR>
A;Cross-references: UNIPROT:Q8ZFA0; GB:AL590842; PIDN:CAC90634.1; PID:G15979840; GSPDB:G
C;Genetics:
A;Gene: YP01817

Query Match      70.4%; Score 38; DB 2; Length 139;
Best Local Similarity 70.0%; Pred. No. 7.4;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VSEGSVEPES 10
   |||||
Db 52 VSEGIPIPEA 61

RESULT 4
T16564
hypothetical protein K05B2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16564
R;Chissole, S.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid K05B2.
A;Reference number: Z18536
A;Accession: T16564
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-808 <CHI>
A;Cross-references: EMBL:U29379; NID:9868192; PID:9868197; PIDN:AAA68732.1; CESP:K05B2.5
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:K05B2.5
A;Introns: 28/2; 40/1; 75/1; 122/3; 241/1; 318/1; 415/3; 500/2; 576/1; 642/1; 684/1; 726/1

Query Match      70.4%; Score 38; DB 2; Length 808;
Best Local Similarity 77.8%; Pred. No. 51;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGSEVPESD 11
   |||||
Db 338 EGSETPKSD 346

RESULT 5
A43817
transforming protein (cbl) - human
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 09-Jul-2004
```

```
C;Accession: A43817
R;Blake, T.J.; Shapiro, M.; Morse III, H.C.; Langdon, W.Y.
Oncogene 6, 653-657, 1991
A;Title: The sequences of the human and mouse c-cbl proto-oncogenes show v-cbl was generated
A;Reference number: A43817; MUID:91232862; PMID:2030914
A;Accession: A43817
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-906 <BLA>
A;Cross-references: UNIPROT:P22681; EMBL:X571110; NID:g29730; PIDN:CAA40393.1; PID:g29731
C;Keywords: DNA binding
F;377-425/Domain: RING finger homology <RRN>

Query Match      70.4%; Score 38; DB 2; Length 906;
Best Local Similarity 77.8%; Pred. No. 57;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVEPVE 9
   |||||
Db 813 VIEGSEQVPE 821

RESULT 6
A83695
hypothetical protein BH0361 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A83695
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83695
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1661 <STO>
A;Cross-references: UNIPROT:Q9KFW0; GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BAB0408
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0361

Query Match      70.4%; Score 38; DB 2; Length 1661;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SEGSEVPESD 11
   |||||
Db 1601 SEDSEVPNSD 1610

RESULT 7
D95284
probable [imported] - Sinorhizobium meliloti (strain 1021) magaplasamid pSymA
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: D95284
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Komp, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: D95284
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <KUR>
A;Cross-references: UNIPROT:Q930L5; GB:AE006469; PIDN:AAK64838.1; PID:g14523251; GSPDB:G
A;Experimental source: strain 1021, megaplasamid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
```

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0335
A;Genome: plasmid
C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 68.5%; Score 37; DB 2; Length 258;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SEGSEVPESD 11
:|||||
Db 103 TDGAEPEDD 112

RESULT 8
S57776
cysteine proteinase (EC 3.4.22.-) - clove pink (fragment)
C;Species: Dianthus caryophyllus (clove pink)
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S57776
P;Jones, M.L.; Larsen, P.B.; Woodson, W.R.
Plant Mol. Biol. 28, 505-512, 1995
A;Title: Ethylene-regulated expression of a carnation cysteine proteinase during flower
A;Reference number: S57776; MUID:95359407; PMID:7632919
A;Accession: S57776
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <JON>
A;Cross-references: UNIPROT:Q43423; EMBL:U17135; NID:G595985; PIDN:AAA79915.1; PID:G5959
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase
P;119,255,276/Active site: Cys, His, Asn #status predicted

Query Match 68.5%; Score 37; DB 2; Length 427;
Best Local Similarity 70.0%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VSEGEVPES 10
:|||||
Db 89 VKEGDELPS 98

RESULT 9
T46146
hypothetical protein T3A5.70 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46146
R;Bloecker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; Sala
submitted to the Protein Sequence Database, December 1999

A;Reference number: 223024
A;Accession: T46146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <ELO>
A;Cross-references: UNIPROT:Q9SCQ7; EMBL:AL132979
A;Experimental source: cultivar Columbia; BAC clone T3A5
C;Genetics:
A;Map position: 3
A;Introns: 369/3
A;Note: T3A5.70

Query Match 68.5%; Score 37; DB 2; Length 447;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SEGSEVPESD 11
:|||||
Db 156 AEGNERPESD 165

RESULT 10
A70363

mannose-1-phosphate guanyltransferase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: A70363
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196686; PMID:9537320
A;Accession: A70363

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-831 <AQF>
A;Cross-references: UNIPROT:O66933; GB:AE000704; NID:G2983301; PIDN:AA06893.1; PID:G2983
A;Experimental source: strain VFS
C;Genetics:
A;Gene: mpg

Query Match 68.5%; Score 37; DB 2; Length 831;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VSEGEVPES 10
:|||||
Db 254 VEEGTPEIPEN 263

RESULT 11
B84683

hypothetical protein At2g28300 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: B84683
R;Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
eusa, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: B84683
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2218 <STO>
A;Cross-references: GB:AE002093; NID:G4803953; PIDN:AA029825.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g28300
A;Map position: 2

Query Match 68.5%; Score 37; DB 2; Length 2218;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VSEGEVPES 10
:|||||
Db 2071 VTEGSELPSS 2080

RESULT 12
E87604

conserved hypothetical protein CC2869 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87604
R;Nierman, W.C.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87604

```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <STO>
A;Cross-references: UNIPROT:Q9A4G5; GB:AE005673; NID:gl3424485; PIDN:AAK24833.1; GSPDB:G
C;Genetics:
A;Gene: CC2869

Query Match          66.7%; Score 36; DB 2; Length 75;
Best Local Similarity 54.5%; Pred. No. 9.2;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VSEGSVEPESD 11
Db 58 LSEGRKIPEPD 68

RESULT 13
T33362
hypothetical protein F16G10.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33362
submitted to the EMBL Data Library, July 1998
R;Gattung, S.; Scheet, P.
A;Description: The sequence of C. elegans cosmid F16G10.
A;Reference number: 221329
A;Accession: T33362
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-204 <GAT>
A;Cross-references: UNIPROT:O76595; EMBL:AF077537; PIDN:AAC26278.1; GSPDB:GNO00020; CESP:
C;Genetics:
A;Experimental source: strain Bristol N2; clone F16G10
A;Gene: CESP:F16G10.10
A;Map position: 2
A;Introns: 56/1; 134/3

Query Match          66.7%; Score 36; DB 2; Length 204;
Best Local Similarity 86.7%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EGSEVPESD 11
Db 120 EGSELPKAD 128

RESULT 14
T52112
deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) [validated] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52112
submitted to the EMBL Data Library, March 1998
R;Ahmad, M.; Jarrillo, J.A.; Cashmore, A.R.
A;Reference number: 225961
A;Accession: T52112
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-496 <AHM>
A;Cross-references: UNIPROT:O24374; EMBL:AF053365; PIDN:AAC08008.1
A;Genetics:
A;Gene: PHR1
A;Function:
A;Description: EC 4.1.99.3 [validated, MUID:97215604]
C;Keywords: carbon-carbon lyase

Query Match          66.7%; Score 36; DB 2; Length 496;
Best Local Similarity 77.8%; Pred. No. 72;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VSEGSVEPVE 9
Db 218 VREGAEVPE 226

```

RESULT 15

```

S51788
malate synthase (EC 4.1.3.2) isoenzyme G - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 23-Aug-1995 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S51788; F65083
R;Molina, I.; Pellicer, M.T.; Badia, J.; Aguilar, J.; Baldona, L.
Eur. J. Biochem. 224, 541-548, 1994
A;Title: Molecular characterization of Escherichia coli malate synthase G. Differentiation
A;Reference number: S51788; MUID:95010032; PMID:7925370
A;Accession: S51788
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-723 <MOL>
A;Cross-references: UNIPROT:P37330; EMBL:X74547; NID:g517246; PIDN:CRA52639.1; PID:g51724
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65083
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-723 <BLAT>
A;Cross-references: GB:AE000380; GB:U00096; NID:gl789344; PIDN:AAC76012.1; PID:gl789348;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: glcB
C;Keywords: carbon-carbon lyase; coenzyme A; oxo-acid-lyase

Query Match          66.7%; Score 36; DB 2; Length 723;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SEGSEVPE 9
Db 352 SEGNEIPE 359

Search completed: September 20, 2005, 12:43:22
Job time : 3.23868 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 5.58751 Seconds
(without alignments)
1008.117 Million cell updates/sec

Title: US-10-671-242-21
Perfect score: 54
Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	236	2 Q96IM4	O6p03 homo sapien
2	54	100.0	285	2 Q6P0R3	O6p03 homo sapien
3	54	100.0	295	2 Q6P1S7	O6p17 homo sapien
4	54	100.0	538	2 Q6SPF0	O6sp0 homo sapien
5	54	100.0	550	2 Q6SPF0	O6sp0 oryctolagus
6	42	77.8	714	2 Q6CC87	O6cc87 yarrowia li
7	41	75.9	244	2 Q27542	O27542 methanobact
8	41	75.9	587	2 Q92VJ3	O92v3 arabisopsis
9	40	74.1	150	2 Q6L129	O6l129 picophilus
10	40	74.1	2972	2 Q8FR52	O8fr52 corynebacte
11	39	72.2	465	2 Q65XA3	O65xa3 oryza sativ
12	38	70.4	139	2 Q8ZFA0	O8zfa0 yersinia pe
13	38	70.4	192	2 Q67LF8	O67lf8 symbiobacte
14	38	70.4	501	1 PTGI RAT	O62969 rattus norv
15	38	70.4	677	2 Q9GYN2	Q9gyn2 caenorhabdi
16	38	70.4	798	2 Q6CUR8	Q6cur8 kluyveromyc
17	38	70.4	906	1 CBL HUMAN	P22681 homo sapien
18	38	70.4	1528	2 Q75B25	Q75b25 ashbya goss
19	38	70.4	1661	2 Q9KFW0	Q9kfw0 bacillus ha
20	38	70.4	1678	2 Q924C5	Q924c5 mus musculu
21	37	68.5	258	2 Q930L5	Q930l5 rhizobium m
22	37	68.5	269	2 Q8L9D7	Q8l9d7 arabisopsis
23	37	68.5	404	2 Q8KE30	Q8ke30 chlorobium
24	37	68.5	427	2 Q43423	Q43423 dianthus ca
25	37	68.5	447	2 Q9SCQ7	Q9scq7 arabisopsis
26	37	68.5	499	2 Q71B18	Q71bf8 stellaria l
27	37	68.5	574	2 Q8NS53	Q8ns53 corynebacte
28	37	68.5	581	2 Q6Y5M8	Q6y5m8 mus musculu
29	37	68.5	675	2 Q7SE16	Q7se16 neurospora
30	37	68.5	768	2 Q64PM0	Q64pm0 bacteroides
31	37	68.5	831	2 Q66933	Q66933 aquifex aeo

32	37	68.5	832	2 Q9N4C1	Q9n4c1 caenorhabdi
33	37	68.5	957	2 Q94GY5	Q94gy5 oryza sativ
34	37	68.5	1512	2 Q8CDL2	Q8cdl2 mus musculu
35	37	68.5	1911	2 Q8OU10	Q8ou10 mus musculu
36	37	68.5	3313	2 Q9W3W6	Q9w3w6 drosophila
37	37	68.5	3571	2 Q9SL27	Q9sl27 arabidopsis
38	37	68.5	3574	2 Q9AUB4	Q9aub4 arabidopsis
39	37	68.5	4685	2 Q93HJ2	Q93hj2 streptomyce
40	36	66.7	75	2 Q9A4G5	Q9a4g5 caulobacter
41	36	66.7	164	2 Q8WMM9	Q8wmm9 ovis aries
42	36	66.7	169	2 Q86TU3	Q86tu3 homo sapien
43	36	66.7	204	2 Q76595	Q76595 caenorhabdi
44	36	66.7	248	2 Q7RZW8	Q7rzw8 neurospora
45	36	66.7	278	1 IOD3_HUMAN	P55073 homo sapien

ALIGNMENTS

RESULT 1
Q96IM4 PRELIMINARY; PRT; 236 AA.
AC Q96IM4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE LOC90378 protein (Hypothetical protein) (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Spletten M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007384; AAH07384.2; -;
DR EMBL; BC080588; AAH80588.1; -;
DR HSSP; P39769; 1KW4.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.

```

DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 54; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 42 VSEGSEVPESD 52
|||||

RESULT 2
Q6POR3 PRELIMINARY; PRT; 285 AA.
AC Q6POR3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Green E.D., Dickinson M.C.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC065477; AAH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 91 VSEGSEVPESD 101
|||||

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC90378 protein (fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Green E.D., Dickinson M.C.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 101 VSEGSEVPESD 111
|||||

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```



```

RN SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY453840; AAR24087.1; -.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM_1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS0105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 54; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 344 VSEGSEVPESD 354
|||||
|||||

RESULT 5
Q6SPE9 PRELIMINARY; PRT; 550 AA.
AC Q6SPE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atherin.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY453841; AAR24088.1; -.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM_1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS0105; SAM_DOMAIN; 1.
SQ SEQUENCE 550 AA; 57019 MW; AF1CF2B780D879A7 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 355 VSEGSEVPESD 365
|||||
|||||

RESULT 6
Q6CC87 PRELIMINARY; PRT; 714 AA.
AC Q6CC87;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome C of strain CLIB99 of Yarrowia
DE lipolytica.
GN ORFNames=YAL10C11495g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

```

```

RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR382129; CAG82035.1; -.
DR GO: GO:0004835; F:ubulin-tyrosine ligase activity; IEA.
DR GO: GO:0006464; P:protein modification; IEA.
DR InterPro: IPR002828; Sure.
DR InterPro: IPR004344; Tub_tyr_lygase.
DR Pfam: PF01975; Sure; 1.
DR Pfam: PF03133; TTL; 1.
DR ProDom: PD005378; Sure; 1.
DR TIGRFAMs: TIGR00087; sure; 1.
SQ SEQUENCE 714 AA; 80159 MW; 26DCA372AD90FE63 CRC64;

Query Match 77.8%; Score 42; DB 2; Length 714;
Best Local Similarity 88.9%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGSEVPESD 11
Db 700 EGSEVPESD 708
|||||
|||||

RESULT 7
O27542 PRELIMINARY; PRT; 244 AA.
AC O27542;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved protein.
GN OrderedLocusNames=MT1498;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Piatrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL: AE000910; BAB85973.1; -.
DR PIR: F69066; F69066.
DR InterPro: IPR010624; KaIC.
DR Pfam: PF06745; KaIC; 1.
KW Complete proteome.
SQ SEQUENCE 244 AA; 27537 MW; A93F15849B5CCC3B CRC64;

```

```

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 150 AA; 17477 MW; 0F81FBE1FC98EA49 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 150;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSGSEVPESD 11
:|||||:
Db 231 ITGEEIPEED 241

RESULT 8
Q9ZVJ3 PRELIMINARY; PRT; 587 AA.
AC Q9ZVJ3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein At2g38720.
GN Name=At2g38720;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005499; AAC67346.1; -.
DR PIR; E84808; E84808.
DR InterPro; IPR007145; MAP65_ASE1.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF03999; MAP65_ASE1; 1.
KW Hypothetical protein.
SQ SEQUENCE 587 AA; 67137 MW; 5A3A70BA72119425 CRC64;

Query Match 75.9%; Score 41; DB 2; Length 587;
Best Local Similarity 81.8%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSGSEVPESD 11
:|||||:
Db 137 VSSGSEVDES 147

RESULT 9
Q6L129 PRELIMINARY; PRT; 150 AA.
AC Q6L129;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PT00738;
OS Picrophilus torridus.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Picrophilaceae; Picrophilus.
OX NCBI_TaxID=82076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 9790 / ATCC 700027;
RX PubMed=15184674; DOI=10.1073/pnas.0401356101;
RA Fuetterer O., Angelov A., Liesegang H., Gottschalk G., Schleper C.,
RA Scheper B., Dock C., Antranikian G., Liebl W.;
RT "Genome sequence of Picrophilus torridus and its implications for life
RT around pH 0."
RL Proc. Natl. Acad. Sci. U.S.A. 101:9091-9096(2004).
DR EMBL; AE017261; AAT43323.1; -.

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 150 AA; 17477 MW; 0F81FBE1FC98EA49 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 150;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSGSEVPE 9
|||||:
Db 36 VSEGSKVPE 44

RESULT 10
Q8FR52 PRELIMINARY; PRT; 2972 AA.
AC Q8FR52;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fatty-acid synthase I (EC 2.3.1.85).
GN Name=fasA; OrderedLocusNames=CE0913;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RA "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP05217; BAC17723.1; -.
DR HSSP; G32472; 11O6.
DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR003965; Fatty_acid_synth.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; Ketoacyl-synt_C; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
KW Acyltransferase; Complete proteome; Transferase.
SQ SEQUENCE 2972 AA; 316363 MW; C0EFF5DDFC123C9 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 2972;
Best Local Similarity 70.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSGSEVPESD 11
:|||||:
Db 2400 AEGTEVPEED 2409

RESULT 11
Q65XA3 PRELIMINARY; PRT; 465 AA.
AC Q65XA3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

```

DE Putative MAP kinase phosphatase.
GN Names:OJ1654_B10.15;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Heu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RT "Oryza sativa BAC OJ1654_B10 genomic sequence.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC108504; AAU44084.1; -;
KW Kinase.
SQ SEQUENCE 465 AA; 50818 MW; E8D433E0006AEF9 CRC64;
Query Match 72.2%; Score 39; DB 2; Length 465;
Best Local Similarity 88.9%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VSEGSEVPE 9
| | | | |
Db 75 VREGSEVPE 83
RESULT 12
Q8ZFAO PRELIMINARY; PRT; 139 AA.
ID Q8ZFAO Q74DW8; O7CHY6;
AC Q8ZFAO; Q74DW8; O7CHY6;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Hypothetical protein YP01817.
GN OrderedLocusNames=Yp1576, YP01817, Y2490;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586350; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Petherston J.C., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,

RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ414150; CAC90634.1; -;
DR EMBL; AB013852; AM86047.1; -;
DR EMBL; AB017132; AAS61810.1; -;
DR PIR; AF0221; AF0221.
DR PFAM; PF03681; UPP0150; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 139 AA; 15549 MW; A608044F830649B1 CRC64;
Query Match 70.4%; Score 38; DB 2; Length 139;
Best Local Similarity 70.0%; Pred. No. 53;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VSEGSEVPE 10
| | | | |
Db 52 VSEGIEPE 61
RESULT 13
Q67LF8 PRELIMINARY; PRT; 192 AA.
ID Q67LF8 Q57LF8;
AC Q67LF8; Q57LF8;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Metalloendopeptidase-related protein.
GN ORFNames=STH2503;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium thermophilum.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005840; BAD41488.1; -;
DR InterPro; IPR011053; Hybrid motif.
DR InterPro; IPR002886; Peptidase_M23B.
DR PFAM; PF01551; Peptidase_M23; 1.
SQ SEQUENCE 192 AA; 20535 MW; C0ABEC13A4F282D0 CRC64;
Query Match 70.4%; Score 38; DB 2; Length 192;
Best Local Similarity 63.6%; Pred. No. 76;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 VSEGSEVPE 11
| | | | |
Db 142 VSEGDRVPOGD 152
RESULT 14
PTGI_RAT STANDARD; PRT; 501 AA.
ID PTGI_RAT
AC Q62969;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Prostacyclin synthase (EC 5.3.99.4) (Prostaglandin 12 synthase).
GN Name=Ptgis; Synonyms=Cyp8;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Geraci M.W., Gao B., Shepherd D., Moore M., Vernon J., Miller Y.E.,
RA Voelkel N.F.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Catalyzes the isomerization of prostaglandin H2 to
 CC prostacyclin (= prostaglandin I2) (By similarity).
 CC -!- CATALYTIC ACTIVITY: (5Z,13E)- (15S)-9-alpha-epidioxy-15-
 CC hydroxyprosta-5,13-dienoate = (5Z,13E)- (15S)-6,9-alpha-epoxy-11-
 CC alpha,15-dihydroxyprosta-5,13-dienoate.
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane. May be
 CC anchored to the membrane via a single transmembrane domain.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U53855; AAB02322.1; -.
 CC DR GCD; 3438; Ptgis.
 CC DR InterPro; IPR001128; Cytochrome_P450.
 CC DR InterPro; IPR002403; EP450IV.
 CC DR Pfam; PF00067; P450; 2.
 CC DR PRINTS; PR00465; EP450IV.
 CC DR PRINTS; PR00385; P450.
 CC DR PROSITE; PS00086; CYTOCHROME P450; FALSE NEG.
 CC DR Endoplasmic reticulum; Heme; Isomerase; Prostaglandin biosynthesis;
 CC Transmembrane.
 CC FT TRANSMEM 1 21 Potential.
 CC FT METAL 442 442 Iron (heme axial ligand) (By similarity).
 CC SQ SEQUENCE 501 AA; 57127 MW; D2D85ADFP3C464863 CRC64;
 CC -----
 CC Query Match 70.4%; Score 38; DB 1; Length 501;
 CC Best Local Similarity 72.7%; Pred. No. 2.3e+02;
 CC Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC ||| : ||| |
 CC ||| : ||| |
 CC QY 1 VSEGEVPEPSD 11
 CC Db 468 VSEDTVEPFD 478
 CC -----
 CC RESULT 15
 CC Q9GYN2
 CC ID Q9GYN2 PRELIMINARY; PRT; 677 AA.
 CC AC Q9GYN2
 CC DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 CC DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 CC DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 CC DE Patterned expression site protein 22.
 CC GN Names=pep-22; ORFNames=K05B2.5;
 CC OS Caenorhabditis elegans.
 CC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC OC Rhabditidae; Peloderinae; Caenorhabditis.
 CC OX NCBI_TaxID=6239;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Bristol N2;
 CC RX MEDLINE=99069613; PubMed=9851916;
 CC RG WormBase Consortium;
 CC RT "Genome sequence of the nematode C. elegans: a platform for
 CC RT investigating biology. The C. elegans Sequencing Consortium. ";
 CC RL Science 282:2012-2018(1998).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Bristol N2;
 CC RA Chisoe S.;
 CC RT "The sequence of C. elegans cosmid K05B2. ";
 CC RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC RN [3]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Bristol N2;
 CC RA Waterston R.;
 CC RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U29379; AAF99981.3; -.
 DR WormBase; WBGene0003986; K05B2.5.
 DR WormPep; K05B2.5; CE34052.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR PROSITE; PS50850; MFS; 1.
 SQ SEQUENCE 677 AA; 73704 MW; 478B6ADE2DF389BE CRC64;
 CC -----
 CC Query Match 70.4%; Score 38; DB 2; Length 677;
 CC Best Local Similarity 77.8%; Pred. No. 3.2e+02;
 CC Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC ||| : ||| |
 CC ||| : ||| |
 CC QY 3 EGSEVPESD 11
 CC Db 269 EGSETPKSD 277
 CC -----
 CC Search completed: September 20, 2005, 12:41:46
 CC Job time : 7.58751 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 6.07222 Seconds
(without alignments)
733.538 Million cell updates/sec

Title: US-10-671-242-21

Perfect score: 54

Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/PCT PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11 NEW PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	11	9	US-09-962-055-21
2	54	100.0	11	9	US-09-976-740-21
3	54	100.0	11	13	US-10-023-529-21
4	54	100.0	11	13	US-10-023-523-21
5	54	100.0	11	15	US-10-616-187-21
6	54	100.0	11	15	US-10-671-242-21
7	54	100.0	12	9	US-09-962-055-27
8	54	100.0	12	9	US-09-976-740-27
9	54	100.0	12	13	US-10-023-529-27
10	54	100.0	12	13	US-10-023-523-27
11	54	100.0	12	15	US-10-616-187-27

12	54	100.0	12	15	US-10-671-242-27
13	54	100.0	26	9	US-09-962-055-20
14	54	100.0	26	9	US-09-976-740-20
15	54	100.0	26	13	US-10-023-529-20
16	54	100.0	26	13	US-10-023-523-20
17	54	100.0	26	15	US-10-616-187-20
18	54	100.0	28	9	US-09-962-055-26
19	54	100.0	28	9	US-09-976-740-26
20	54	100.0	28	13	US-10-023-529-26
21	54	100.0	28	13	US-10-023-523-26
22	54	100.0	28	15	US-10-616-187-26
23	54	100.0	28	15	US-10-671-242-26
24	54	100.0	217	9	US-09-962-055-7
25	54	100.0	217	9	US-09-976-740-7
26	54	100.0	217	13	US-10-023-529-7
27	54	100.0	217	13	US-10-023-523-7
28	54	100.0	217	15	US-10-616-187-7
29	54	100.0	217	15	US-10-671-242-7
30	54	100.0	232	9	US-09-962-055-3
31	54	100.0	232	9	US-09-976-740-3
32	54	100.0	232	13	US-10-023-529-3
33	54	100.0	232	13	US-10-023-523-3
34	54	100.0	232	15	US-10-616-187-3
35	54	100.0	232	15	US-10-671-242-3
36	54	100.0	241	9	US-09-925-298-665
37	54	100.0	241	14	US-10-102-806-665
38	54	100.0	252	9	US-09-962-055-4
39	54	100.0	252	9	US-09-976-740-4
40	54	100.0	252	13	US-10-023-529-4
41	54	100.0	252	13	US-10-023-523-4
42	54	100.0	252	15	US-10-616-187-4
43	54	100.0	252	15	US-10-671-242-4
44	54	100.0	317	9	US-09-962-055-2
45	54	100.0			

ALIGNMENTS

RESULT 1

US-09-962-055-21

; Sequence 21, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-962-055-21

Query Match          100.0%; Score 54; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
Db 1 VSEGSSEVPESD 11

RESULT 2
US-09-976-740-21
; Sequence 21, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-21

Query Match          100.0%; Score 54; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
Db 1 VSEGSSEVPESD 11

RESULT 3
US-10-023-529-21
; Sequence 21, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-21

Query Match          100.0%; Score 54; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
Db 1 VSEGSSEVPESD 11

RESULT 4
US-10-023-523-21
; Sequence 21, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-21

Query Match          100.0%; Score 54; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
Db 1 VSEGSSEVPESD 11
```

RESULT 5
US-10-616-187-21
; Sequence 21, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-21

Query Match 100.0%; Score 54; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 6
US-10-671-242-21
; Sequence 21, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-21

Query Match 100.0%; Score 54; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 7
US-09-962-055-27
; Sequence 27, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-962-055-27

Query Match 100.0%; Score 54; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 2 VSEGSVPESD 12

RESULT 8
US-09-976-740-27

```

; Sequence 27, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-976-740-27

```

```

Query Match      100.0%; Score 54; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 VSEGSVPESD 11
         |||||
Db       2 VSEGSVPESD 12

```

RESULT 9

```

US-10-023-529-27
; Sequence 27, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-529-27

```

```

Query Match      100.0%; Score 54; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;

```

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSVPESD 11
         |||||
Db       2 VSEGSVPESD 12

```

RESULT 10

```

US-10-023-523-27
; Sequence 27, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-523-27

```

```

Query Match      100.0%; Score 54; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 VSEGSVPESD 11
         |||||
Db       2 VSEGSVPESD 12

```

RESULT 11

```

US-10-616-187-27
; Sequence 27, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547

```



```

; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-616-187-27

Query Match      100.0%; Score 54; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSEVPESD 11
   |||||
Db 2 VSEGSEVPESD 12

RESULT 12
US-10-671-242-27
; Sequence 27, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671.242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-671-242-27

Query Match      100.0%; Score 54; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSEVPESD 11
   |||||
Db 2 VSEGSEVPESD 12

RESULT 13
US-09-962-055-20
; Sequence 20, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 54; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSEVPESD 11
   |||||
Db 16 VSEGSEVPESD 26

RESULT 14
US-09-976-740-20
; Sequence 20, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT

```

```
; ORGANISM: Homo sapiens
US-09-976-740-20

Query Match      100.0%; Score 54; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSEVPESD 11
Db      16 VSEGSEVPESD 26

RESULT 15
US-10-023-529-20
; Sequence 20, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 54; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSEVPESD 11
Db      16 VSEGSEVPESD 26

Search completed: September 20, 2005, 12:53:06
Job time : 6.07222 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 5.94859 Seconds
(without alignments)
650.171 Million cell updates/sec

Title: US-10-671-242-22

Perfect score: 58

Sequence: 1 EDDPPDGLG 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	58	100.0	10	4	AAB82812 Human low
2	58	100.0	132	7	AD59158 Human Pro
3	58	100.0	217	2	AAW49041 Human low
4	58	100.0	217	4	AAB82803 Human low
5	58	100.0	241	3	AAB88957 Human low
6	58	100.0	538	4	AAB82806 Human low
7	54	93.1	232	4	AAB82799 Rabbit lo
8	54	93.1	252	4	AAB82800 Rabbit lo
9	54	93.1	317	2	AAW49038 Rabbit lo
10	54	93.1	317	4	AAB82798 Rabbit lo
11	54	93.1	550	4	AAB82807 Rabbit lo
12	44	75.9	945	4	AAW51653 Arabidops
13	43	74.1	100	3	AAW02194 Human sec
14	43	74.1	346	5	ABB06721 Human wit
15	43	74.1	804	4	ABB89136 Drosophili
16	41	70.7	140	5	ABP31168 Human ORF
17	41	70.7	354	4	AAU41146 Propionib
18	41	70.7	354	6	ABM40665 Propionib
19	41	70.7	354	7	AD31306 Human nov
20	41	70.7	354	8	ADR46143 Human pro
21	41	70.7	354	8	ADR46142 Human pro
22	41	70.7	435	8	ADQ65987 Novel hum
23	41	70.7	510	8	ADS29627 Bacterial
24	39	67.2	14	8	ADI00477 Influenza
25	39	67.2	14	8	ADI56752 Flu virus

ALIGNMENTS

RESULT 1

AAB82812
ID AAB82812 standard; peptide; 10 AA.

XX AC AAB82812;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;

KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US0063556.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

PI Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

DR New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

PT Claim 14(a); Page: 143pp; English.

PS The present sequence is that of a peptide comprising amino acid residues 529-538 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs.

CC Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for

CC atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine

Abp02064 Human ORF
Abu44831 Protein e
Adh39670 Streptomy
Adh39812 Streptomy
Adn20818 Bacterial
Adal9453 Mouse LAT
Aam25741 Human pro
Aae06606 Human pro
Ade07899 Novel pro
Aam94615 Human rep
Abm84853 Human dia
Aay01459 Polypepti
Abo23396 Amino aci
Adr20783 Human sec
Aay41706 Human PRO
Aay01458 Polypepti
Aab33463 Human PRO
Aay94975 Human sec
Aab44262 Human PRO
Aab24021 Human PRO

CC compositions. Note: the present sequence is not shown in the
CC specification but is derived from the human LBP-2 sequence given in
CC figure 7A (see AAS82806)

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
|||||
Db 1 EDDDPDGFGLG 10

RESULT 2
ID ADE59158
XX ADE59158 standard; protein; 132 AA.
AC ADE59158;
XX
DT 29-JAN-2004 (first entry)
XX Human Protein XP_031299, SEQ ID NO 5049.
DE
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX WO2003016475-A2.
PN
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-033347P.
XX
XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; XP_031299.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 132 AA;

Query Match 100.0%; Score 58; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
|||||
Db 123 EDDDPDGFGLG 132

RESULT 3
ID AAW49041
XX AAW49041 standard; protein; 217 AA.
AC AAW49041;
XX
DT 09-NOV-1998 (first entry)
XX
XX Human low density lipoprotein binding protein LBP-2.
DE
XX Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH FT Peptide 8..33 /note= "Claim 2"
FT FT Peptide 8..22 /note= "Claim 2"
FT FT Peptide 23..33 /note= "Claim 2"
FT FT Peptide 208..217 /note= "Claim 2"
FT
XX WO9823282-A1.
PN
XX
XX 04-JUN-1998.
PD
XX 26-NOV-1997; 97WO-US021857.
PF
XX 27-NOV-1996; 96US-0031930P.
PR 03-JUN-1997; 97US-0048547P.
XX
XX (BOST-) BOSTON HEART FOUND INC.
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
PI WPI; 1998-322455/28.
DR N-PSDB; AAV32838.
XX
XX Nucleic acid encoding low density lipoprotein binding proteins and
PT related vectors - transformed cells, proteins, and modulators of binding,
PT useful for treatment, and diagnosis of atherosclerosis and for identifying
PT subjects at risk.
XX
XX Claim 1; Fig 7; 47pp; English.

CC This polypeptide comprises novel human low density lipoprotein (LDL)
CC binding protein LBP-2 that is capable of binding both native and methyl
CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human
CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP

CC Metabolism or structure is diagnostic of a risk for atherosclerosis. The
 CC invention provides: methods for determining if an animal is at risk for
 CC atherosclerosis (e.g. for prenatal screening); methods for treating
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
 CC methods for treating a cell having an abnormality in LBP structure or
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
 CC well as recombinant vectors and host cells used to produce recombinant
 CC LBP

SQ Sequence 217 AA;

Query Match 100.0%; Score 58; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
 |||||
 Db 208 EDDDPDGFGLG 217

RESULT 4

AAB82803
 ID AAB82803 standard; protein; 217 AA.

AC AAB82803;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2).

KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

OS Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAH26494.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13 (g); Fig 7B; 143pp; English.

XX The present sequence is that of the N-terminal portion of novel human low
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
 CC deduced from an isolated partial cDNA clone (see AAH26494). A full-length
 CC sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP
 CC polypeptides of the invention that are capable of binding to native and
 CC methylated low density lipoproteins. Also claimed are biologically active
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
 CC as expression vectors, cells and methods of producing the LBPs. Methods
 CC of determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed

XX Sequence 217 AA;

Query Match 100.0%; Score 58; DB 4; Length 217;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
 |||||
 Db 208 EDDDPDGFGLG 217

RESULT 5

AAB58957

ID AAB58957 standard; protein; 241 AA.

XX AAB58957;

DT 27-MAR-2001 (first entry)

XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.

KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neutropenic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; anticancer; vulnary; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005881.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX N-PSDB; AAF21860.

XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 PT neurological diseases.

XX Claim 11; Page 1112; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neutropenic;
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 CC antiinflammatory; anticancer; vulnary; anticonvulsant; antibacterial;
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases

XX Sequence 241 AA;

Query Match 100.0%; Score 58; DB 3; Length 241;
 Best Local Similarity 100.0%; Pred. No. 0.14; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
 |||||
 Db 232 EDDDPDGFLG 241

RESULT 6
 AAB82806
 ID AAB82806 standard; protein; 538 AA.
 XX
 AC AAB82806;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Human low density lipoprotein binding protein 2 (LBP-2).
 XX
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US006356.
 PR 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 DR WPI; 2001-565505/63.
 DR N-PSDB; AAH26499.
 XX
 PT New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX
 PS Claim 13(j); Fig 7A; 143pp; English.
 XX

The present sequence is that of novel human low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence was deduced from the coding region of isolated genomic DNA (see AAH26499). It differs from the sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the presence of an additional 321 amino acids at the N-terminus (the cDNA clone is 5' truncated). Human LBP-2 is an example of claimed LBP methyleated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529 CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of CC determining if an animal is at risk for atherosclerosis, methods for CC evaluating an agent for use in treating atherosclerosis, and methods for CC treating a cell having an abnormality in structure or metabolism of LBP CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or CC nucleic acid, and vaccine compositions, are also claimed

Sequence 538 AA;
 Query Match 100.0%; Score 58; DB 4; Length 538;
 Best Local Similarity 100.0%; Pred. No. 0.31; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
 |||||
 Db 529 EDDDPDGFLG 538

RESULT 7
 AAB82799
 ID AAB82799 standard; protein; 232 AA.
 XX
 AC AAB82799;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
 XX
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US006356.
 PR 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 DR WPI; 2001-565505/63.
 DR N-PSDB; AAH26489.
 XX
 PT New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX
 PS Claim 13(c); Fig 3; 143pp; English.
 XX

The present sequence is that of a partial sequence of novel rabbit low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is deduced from an isolated cDNA clone (see AAH26489). Full-length rabbit LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed polypeptides of the invention, termed LBPs, that are capable of binding to native and methylated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine compositions

Sequence 232 AA;
 Query Match 93.1%; Score 54; DB 4; Length 232;
 Best Local Similarity 90.0%; Pred. No. 0.61; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
 |||||
 Db 223 EDDDPDGFLG 232

RESULT 8
 AAB82800
 ID AAB82800 standard; protein; 252 AA.
 XX
 AC AAB82800;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 3 (LBP-3).

XX Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX Oryctolagus cuniculus.
 XX W0200164874-A2.
 XX 07-SEP-2001.
 XX 28-FEB-2001; 2001WO-US006356.
 XX 02-MAR-2000; 2000US-00517849.
 XX 14-JUL-2000; 2000US-00616289.
 XX (BOST-) BOSTON HEART FOUND INC.
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 2001-565505/63.
 XX N-PSDB; AAH26490.
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX Claim 13(d); Fig 4; 143pp; English.
 XX The present sequence is that of a partial sequence of novel rabbit low
 CC density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is
 CC deduced from an isolated cDNA clone (see AAH26490). A full-length
 CC sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an
 CC example of claimed polypeptides of the invention, termed LBPs, that are
 CC capable of binding to native and methylated low density lipoproteins.
 CC Also claimed are biologically active fragments and analogues of LBPs.
 CC polynucleotides encoding LBPs, as well as expression vectors, cells, and
 CC methods of producing the LBPs. Methods for determining if an animal is at
 CC risk for atherosclerosis, methods for evaluating an agent for use in
 CC treating atherosclerosis, and methods for treating a cell having an
 CC abnormality in structure or metabolism of LBP are also claimed, as are
 CC pharmaceutical compositions comprising an LBP polypeptide or nucleic
 CC acid, and vaccine compositions
 XX Sequence 252 AA;
 SQ
 Query Match 93.1%; Score 54; DB 4; Length 252;
 Best Local Similarity 90.0%; Pred. No. 0.67;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDDDPDGFLG 10
 Db 243 EDDDPDGFLG 252
 RESULT 9
 AAH49038
 ID AAH49038 standard; protein; 317 AA.
 XX
 AC AAH49038;
 XX
 DT 09-NOV-1998 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein LBP-2.
 DE
 KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
 KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.
 XX Oryctolagus cuniculus.
 XX
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 10
 FT /note= "encoded by TAG"
 FT Protein 66..317
 FT /note= "Claim 1"

FT Protein 86..317
 FT /note= "Claim 1"
 FT Peptide 105..132
 FT /note= "Claim 2"
 FT Peptide 105..120
 FT /note= "Claim 2"
 FT Peptide 121..132
 FT /note= "Claim 2"
 FT Peptide 211..220
 FT /note= "Claim 2"
 XX W09823282-A1.
 PN 04-JUN-1998.
 XX 26-NOV-1997; 97WO-US021857.
 XX 27-NOV-1996; 96US-0031930P.
 PR 03-JUN-1997; 97US-0048547P.
 XX (BOST-) BOSTON HEART FOUND INC.
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 1998-322455/28.
 XX N-PSDB; AAV32835.
 XX Nucleic acid encoding low density lipoprotein binding proteins and
 PT related vectors - transformed cells, proteins, and modulators of binding,
 PT useful for treatment and diagnosis of atherosclerosis and for identifying
 PT subjects at risk.
 XX Claim 1; Fig 2; 47pp; English.
 XX This polypeptide comprises novel rabbit low density lipoprotein (LDL)
 CC binding protein LBP-2 that is capable of binding both native and methyl
 CC LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA
 CC (see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and
 CC human LBPs (see AAH49037-42) are claimed. An abnormality in an aspect of
 CC LBP metabolism or structure is diagnostic of a risk for atherosclerosis.
 CC The invention provides: methods for determining if an animal is at risk
 CC for atherosclerosis (e.g. for prenatal screening); methods for treating
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
 CC methods for treating a cell having an abnormality in LBP structure or
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
 CC well as recombinant vectors and host cells used to produce recombinant
 CC LBP
 XX
 SQ Sequence 317 AA;
 Query Match 93.1%; Score 54; DB 2; Length 317;
 Best Local Similarity 90.0%; Pred. No. 0.84;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDDDPDGFLG 10
 Db 308 EDDDPDGFLG 317
 RESULT 10
 AAB82798
 ID AAB82798 standard; protein; 317 AA.
 XX
 AC AAB82798;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
 DE
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
 XX

```

OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "encoded by TAG"
XX
PN WO200164874-A2.
XX
PD 07-SEP-2001.
XX
XX
PF 28-FEB-2001; 2001WO-US006356.
XX
PR 02-MAR-2000; 2000US-00517849.
PR 14-JUL-2000; 2000US-00616289.
XX
PA (BOST-) BOSTON HEART FOUND INC.
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX N-PSDB; AAH26488.
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
XX Claim 13(b); Fig 2b; 143pp; English.
XX
XX The present sequence is that of a partial sequence of novel rabbit low
density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
deduced from an isolated cDNA clone (see AAH26488). Full-length rabbit
LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed
polypeptides of the invention, termed LBPs, that are capable of binding
to native and methylated low density lipoproteins. Also claimed are
biologically active fragments and analogues of LBPs, polynucleotides
encoding LBPs, as well as expression vectors, cells and methods of
producing the LBPs. Methods of determining if an animal is at risk for
atherosclerosis, methods for evaluating an agent for use in treating
atherosclerosis, and methods for treating a cell having an abnormality in
structure or metabolism of LBP are also claimed, as are pharmaceutical
compositions comprising an LBP polypeptide or nucleic acid, and vaccine
compositions
XX
SQ Sequence 317 AA;

Query Match 93.1%; Score 54; DB 4; Length 317;
Best Local Similarity 90.0%; Pred. No. 0.84;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
DB 308 EDDDDPGFLG 317
|||||:|||||

RESULT 11
AAB82807
ID AAB82807 standard; protein; 550 AA.
XX
AC AAB82807;
XX
XX 12-NOV-2001 (first entry)
XX
DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
XX
KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
XX atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX
OS Oryctolagus cuniculus.
XX
XX WO200164874-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US006356.
XX
PF Lindquist S, Queitsch C, Vierling E;
XX WPI; 2001-639123/73.
XX

XX Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "encoded by TAG"
XX
PN WO200164874-A2.
XX
PD 07-SEP-2001.
XX
XX
PF 28-FEB-2001; 2001WO-US006356.
XX
PR 02-MAR-2000; 2000US-00517849.
PR 14-JUL-2000; 2000US-00616289.
XX
PA (BOST-) BOSTON HEART FOUND INC.
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX N-PSDB; AAH26500.
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
XX Claim 13(k); Fig 2a; 143pp; English.
XX
XX The present sequence is that of a full-length sequence of novel rabbit
low density lipoprotein binding protein 2 (LBP-2). The amino acid
sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit
LBP-2 is an example of claimed polypeptides of the invention, termed
LBPs, that are capable of binding to native and methylated low density
lipoproteins. Also claimed are biologically active fragments and
analogues of LBPs, polynucleotides encoding LBPs, as well as expression
vectors, cells and methods of producing the LBPs. Polypeptides having
amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-
18) of the present sequence are claimed. Methods of determining if an
animal is at risk for atherosclerosis, methods for evaluating an agent
for use in treating atherosclerosis, and methods for treating a cell
having an abnormality in structure or metabolism of LBP are also claimed,
as are pharmaceutical compositions comprising an LBP polypeptide or
nucleic acid, and vaccine compositions
XX
SQ Sequence 550 AA;

Query Match 93.1%; Score 54; DB 4; Length 550;
Best Local Similarity 90.0%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
DB 541 EDDDDPGFLG 550
|||||:|||||

RESULT 12
AAM51653
ID AAM51653 standard; protein; 945 AA.
XX
AC AAM51653;
XX
XX 14-JAN-2002 (first entry)
XX
DE Arabidopsis thaliana ERD1 protein (GenBank: BAA04506).
XX
KW Transgenic plant; stress tolerance; heat shock protein; HSP; cotton;
XX canola; soybean; corn; wheat; tobacco; sorghum; potato; tomato;
XX Arabidopsis thaliana.
XX
OS Arabidopsis thaliana.
XX
XX WO200170929-A2.
XX
XX 27-SEP-2001.
XX
XX 20-MAR-2001; 2001WO-US008836.
XX
XX 20-MAR-2000; 2000US-0190769P.
XX
XX 18-APR-2000; 2000US-0198116P.
XX
XX (ARCH-) ARCH DEV CORP.
XX
XX Lindquist S, Queitsch C, Vierling E;
XX WPI; 2001-639123/73.
XX

```


DR N-PSDB; AA166059.
 XX Transgenic plants with improved heat stress tolerance, useful for
 PT producing animal feed, oil and synthetic products.
 XX
 XX
 PS Claim 2; Page; 91pp; English.
 XX
 CC The invention relates to a transgenic plant, comprising a genetic
 CC construct comprising a promoter operatively linked to a nucleic acid
 CC sequence (AA166057-AA166084) encoding a plant Heat Shock Protein (HSP)
 CC family amino acid sequence (AA166057-AA166084). The transgenic plant has
 CC increased stress tolerance, especially to heat. The plant is a cereal,
 CC grass, ornamental plant, crop plant, food plant, oil-producing plant, a
 CC synthetic product-producing plant, an environmental waste absorbing
 CC plant, an alcohol plant, a medicinal plant, a recreational plant and/or
 CC an animal feed plant. In particular, the transgenic plant is cotton,
 CC canola, soybean, corn, wheat, tobacco, sorghum, potato, tomato and/or
 CC Arabidopsis thaliana. The plants may be used to produce animal feed,
 CC alcohol, crop, oil, medicine or a synthetic product. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained from GenBank using the Accession Number reference provided
 CC in the specification
 XX
 SQ Sequence 945 AA;
 Query Match 75.9%; Score 44; DB 4; Length 945;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EDDPDGFLG 10
 Db 118 EDRDPQFLG 127
 |||||
 |||||
 RESULT 13
 AAG02194
 ID AAG02194 standard; protein; 100 AA.
 XX
 AC AAG02194;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 6275.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GIST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC02200.
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 13; SEQ ID NO 6275; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different

CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 XX
 SQ Sequence 100 AA;
 Query Match 74.1%; Score 43; DB 3; Length 100;
 Best Local Similarity 70.0%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EDDPDGFLG 10
 Db 83 DDDDDGFFG 92
 :|||
 :|||
 RESULT 14
 ABB06721
 ID ABB06721 standard; protein; 346 AA.
 XX
 AC ABB06721;
 XX
 DT 12-JUN-2002 (first entry)
 XX
 DE Human wither inhibitor 38.06 protein SEQ ID NO:2.
 DE Human; wither inhibitor 38.06; inhibitor; viral infection; tumour.
 KW
 XX Homo sapiens.
 OS
 PN CN1328021-A.
 XX
 PD 26-DEC-2001.
 XX
 PF 12-JUN-2000; 2000CN-00116421.
 XX
 PR 12-JUN-2000; 2000CN-00116421.
 XX
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI; 2002-242307/30.
 DR N-PSDB; ABL50160.
 XX
 PT Polypeptide-human wither inhibitor 38.06 and polynucleotide for coding
 PT it.
 XX
 PS Claim 1; Page 27-28 (Disclosure); 34pp; Chinese.
 XX
 CC The present sequence represents human wither inhibitor 38.06 (I). The
 CC present invention also describes a process for preparing (i) using DNA
 CC recombination techniques. (i) and the polynucleotide encoding it can be
 CC used in the treatment of diseases such as viral infection diseases and
 CC tumours
 CC
 SQ Sequence 346 AA;
 Query Match 74.1%; Score 43; DB 5; Length 346;
 Best Local Similarity 70.0%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EDDPDGFLG 10
 Db 82 DDDDDGFFG 91
 :|||
 :|||

RESULT 15
ABB69136
ID ABB69136 standard; protein; 804 AA.
XX
AC ABB69136;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 34200.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL13239.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 34200; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 804 AA;

Query Match 74.1%; Score 43; DB 4; Length 804;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 DDDPDGFLG 10
Db 483 DDSPEGLG 491

Search completed: September 20, 2005, 12:34:47
Job time : 7.94859 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 1.51775 Seconds
(without alignments)
491.841 Million cell updates/sec

Title: US-10-671-242-22
Perfect score: 58
Sequence: 1 EDDDPGFLG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	10	3	US-08-979-608A-22
2	58	100.0	10	4	US-09-517-849-22
3	58	100.0	10	4	US-09-616-289-22
4	58	100.0	217	3	US-08-979-608A-7
5	58	100.0	217	4	US-09-517-849-7
6	58	100.0	217	4	US-09-616-289-7
7	58	100.0	538	4	US-09-616-289-43
8	54	93.1	232	3	US-08-979-608A-3
9	54	93.1	232	4	US-09-517-849-3
10	54	93.1	232	4	US-09-616-289-3
11	54	93.1	252	3	US-08-979-608A-4
12	54	93.1	252	4	US-09-517-849-4
13	54	93.1	252	4	US-09-616-289-4
14	54	93.1	317	3	US-08-979-608A-2
15	54	93.1	317	4	US-09-517-849-2
16	54	93.1	317	4	US-09-616-289-2
17	54	93.1	550	4	US-09-616-289-47
18	43	74.1	100	4	US-09-513-999C-6275
19	41	70.7	109	4	US-09-621-976-7055
20	38	65.5	328	4	US-09-902-540-13561
21	38	65.5	826	4	US-09-248-796A-14387
22	37	63.8	198	3	US-09-228-986-93
23	37	63.8	198	4	US-10-101-464A-93
24	37	63.8	211	4	US-10-101-464A-762
25	37	63.8	227	4	US-09-248-796A-21586
26	37	63.8	260	4	US-09-589-927-8
27	37	63.8	260	4	US-09-277-665-8

28	37	63.8	260	4	US-09-589-987-8	Sequence 8, Appli
29	37	63.8	455	2	US-08-738-172-4	Sequence 4, Appli
30	37	63.8	485	4	US-09-252-991A-17170	Sequence 17170, A
31	37	63.8	575	4	US-09-248-796A-14348	Sequence 14348, A
32	37	63.8	596	4	US-10-101-464A-889	Sequence 889, App
33	37	63.8	996	4	US-10-101-464A-933	Sequence 933, App
34	36	62.1	263	4	US-09-902-540-12633	Sequence 12633, A
35	36	62.1	354	4	US-09-949-016-10579	Sequence 10579, A
36	36	62.1	427	4	US-09-248-796A-19520	Sequence 19520, A
37	36	62.1	459	4	US-09-252-991A-24552	Sequence 24552, A
38	36	62.1	469	4	US-09-411-132A-5	Sequence 5, Appli
39	36	62.1	703	4	US-09-252-991A-20065	Sequence 20065, A
40	36	62.1	704	1	US-08-190-802A-62	Sequence 62, Appli
41	36	62.1	704	1	US-08-188-583-5	Sequence 5, Appli
42	36	62.1	704	1	US-08-646-715-5	Sequence 5, Appli
43	36	62.1	704	2	US-08-308-818-3	Sequence 3, Appli
44	36	62.1	704	3	US-08-477-346-62	Sequence 62, Appli
45	36	62.1	704	3	US-08-473-089-62	Sequence 62, Appli

ALIGNMENTS

RESULT 1

US-08-979-608A-22

; Sequence 22, Application US/08979608A

; Patent No. 6355451

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M. S.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/979,608A

; FILING DATE: 26-NOV-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,547

; FILING DATE: 03-JUN-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 22:

US-08-979-608A-22

Query Match 100.0%; Score 58; DB 3; Length 10;

```
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 1 EDDDPDGFGLG 10

RESULT 2
US-09-517-849-22
; Sequence 22, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-517-849-22

Query Match 100.0%; Score 58; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 1 EDDDPDGFGLG 10

RESULT 3
US-09-616-289-22
; Sequence 22, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 1 EDDDPDGFGLG 10

RESULT 4
US-08-979-608A-7
; Sequence 7, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 217 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match 100.0%; Score 58; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDDPDGFGLG 10
Db 208 EDDDPDGFGLG 217

RESULT 5

US-09-517-849-7
Sequence 7, Application US/09517849
Patent No. 6605588
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-Nov-1997

ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match 100.0%; Score 58; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDDPDGFGLG 10
Db 208 EDDDPDGFGLG 217

RESULT 6

Query Match 100.0%; Score 58; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-616-289-7
Sequence 7, Application US/09616289
Patent No. 6632923
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 217
TYPE: PRT
ORGANISM: Homo sapiens
US-09-616-289-7

Query Match 100.0%; Score 58; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDDPDGFGLG 10
Db 208 EDDDPDGFGLG 217

RESULT 7

US-09-616-289-43
Sequence 43, Application US/09616289
Patent No. 6632923
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 538
TYPE: PRT
ORGANISM: Homo sapiens
US-09-616-289-43

```
QY      1 EDDDPDGLG 10
Db      529 EDDDPDGLG 538

RESULT 8
US-08-979-608A-3
; Sequence 3, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;           Lees, Robert S.
;           Law, Simon W.
;           Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

QY      1 EDDDPDGLG 10
Db      529 EDDDPDGLG 538

Query Match      93.1%; Score 54; DB 3; Length 232;
Best Local Similarity 90.0%; Pred. No. 0.08;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
US-09-517-849-3
; Sequence 3, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;           Lees, Robert S.
;           Law, Simon W.
;           Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3

QY      1 EDDDPDGLG 10
Db      223 EDDDPDGLG 232

Query Match      93.1%; Score 54; DB 4; Length 232;
Best Local Similarity 90.0%; Pred. No. 0.08;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
US-09-616-289-3
; Sequence 3, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
;           Lees, Ann M.
;           Law, Simon W.
;           Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
```

```
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-3

Query Match          93.1%; Score 54; DB 4; Length 232;
Best Local Similarity 90.0%; Pred. No. 0.08;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
    |||||:||||
Db 223 EDDDPDGFLG 232

RESULT 11
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 635451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match          93.1%; Score 54; DB 3; Length 252;
Best Local Similarity 90.0%; Pred. No. 0.088;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
    |||||:||||
Db 243 EDDDPDGFLG 252

RESULT 12
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 635451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match          93.1%; Score 54; DB 3; Length 252;
Best Local Similarity 90.0%; Pred. No. 0.088;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
    |||||:||||
Db 243 EDDDPDGFLG 252

RESULT 13
US-09-616-289-4
; Sequence 4, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
```

```
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-4

Query Match          93.1%; Score 54; DB 4; Length 252;
Best Local Similarity 90.0%; Pred. No. 0.088;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 243 EDDDPDGFLG 252

RESULT 14
US-08-979-608A-2
; Sequence 2, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-979-608A-2

Query Match          93.1%; Score 54; DB 3; Length 317;
Best Local Similarity 90.0%; Pred. No. 0.11;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 308 EDDDPDGFLG 317

RESULT 15
US-09-517-849-2
; Sequence 2, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 02-Mar-2000
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-517-849-2

Query Match          93.1%; Score 54; DB 4; Length 317;
Best Local Similarity 90.0%; Pred. No. 0.11;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 308 EDDDPDGFLG 317

Search completed: September 20, 2005, 12:45:31
Job time : 2.51775 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 1.12607 Seconds
(without alignments)
854.447 Million cell updates/sec

Title: US-10-671-242-22

Perfect score: 58

Sequence: 1 EDDDPDGFLLG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	75.9	406	1 A31050	calsequestrin prec
2	44	75.9	945	2 JN0901	calsequestrin prec
3	41	70.7	420	1 S22418	calsequestrin prec
4	41	70.7	510	2 E84347	glycerol kinase [i
5	39	67.2	539	2 D71260	hypothetical prote
6	38	65.5	232	2 G75608	hypothetical prote
7	38	65.5	472	2 T45219	regulator of pre-m
8	38	65.5	526	2 T51476	hypothetical prote
9	38	65.5	784	2 D84386	ATP-dependent RNA
10	38	65.5	1047	2 T49425	hypothetical prote
11	37	63.8	127	2 AH3491	hypothetical cytos
12	37	63.8	186	2 AD3306	periplasmic dipept
13	37	63.8	228	2 AG3480	25K outer-membrane
14	37	63.8	243	2 H90015	hypothetical prote
15	37	63.8	380	1 S74313	sensory transducti
16	37	63.8	428	2 AH07093	succinylglutamic s
17	37	63.8	450	2 E82973	conserved hypothet
18	37	63.8	516	1 DACHA	procollagen-prolin
19	37	63.8	533	2 B56110	tyrosine phosphopr
20	37	63.8	584	2 S51882	topoisomerase I-re
21	37	63.8	953	2 T21292	hypothetical prote
22	37	63.8	1121	2 T52631	1-phosphatidylinos
23	36	62.1	111	2 T04405	phytochrome C - ba
24	36	62.1	151	2 T33918	hypothetical prote
25	36	62.1	153	2 A41510	myosin regulatory
26	36	62.1	225	2 AC1401	adenylate kinases
27	36	62.1	236	2 AI3377	25K outer-membrane
28	36	62.1	265	2 T46173	hypothetical prote
29	36	62.1	271	2 T42311	hypothetical prote

RESULT 1

A31050

calsequestrin precursor, cardiac and skeletal muscle - chicken
N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-1
C;Species: Gallus gallus (Chicken)
C;Date: 16-Jul-1999 #sequence revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: A31050; B31049; A35709; A34652
R;Clegg, D.O.; Helder, J.C.; Hann, B.C.; Hall, D.E.; Reichardt, L.F.
J. Cell Biol. 107, 699-705, 1988
A;Title: Amino acid sequence and distribution of mRNA encoding a major skeletal muscle la
domain.
A;Reference number: A31050; MUID:88331074; PMID:3417769

ALIGNMENTS

30	36	62.1	349	2	T36152	probable transpos
31	36	62.1	390	1	A60424	calsequestrin prec
32	36	62.1	393	2	T33917	hypothetical prote
33	36	62.1	397	2	E69173	protoporphyrin IX
34	36	62.1	407	2	A10595	imidazolonepropion
35	36	62.1	421	2	G98730	hypothetical prote
36	36	62.1	469	2	B43357	pancreatic lipase-
37	36	62.1	470	2	A54232	lipase, CoPL-RP2 -
38	36	62.1	474	2	G84543	probable protein k
39	36	62.1	482	1	A46696	triacylglycerol li
40	36	62.1	482	1	A34671	triacylglycerol li
41	36	62.1	544	2	T29325	hypothetical prote
42	36	62.1	569	2	T47358	hypothetical prote
43	36	62.1	701	2	JN0674	ubiquitin-like fus
44	36	62.1	704	2	S33263	transcription init
45	36	62.1	821	1	B34488	calpain (EC 3.4.22

A;Reference number: A34652; MUID:90147805; PMID:2302244
A;Accession: A34652
A;Status: preliminary
A;Molecule type: protein
A;Residues: 20-39 <YAZ>
R;Yazaki, P.J.; Salvatori, S.; Dahms, A.S.
Biochem. Biophys. Res. Commun. 170, 1089-1095, 1990
A;Title: Amino acid sequence of chicken calsequestrin deduced from c DNA: comparison of c
A;Reference number: A35709; MUID:90358803; PMID:2390076
A;Accession: A35709
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 20-147, '1', 149-406 <YAZ>
A;Cross-references: GB:M58048; NID:G211496; PIDN:AAA48674.1; PID:G211497
R;Yazaki, P.J.; Salvatori, S.; Sabbadini, R.A.; Dahms, A.S.
Biochem. Biophys. Res. Commun. 166, 898-903, 1990
A;Title: Calsequestrin, and intracellular calcium-binding protein of skeletal muscle sarco

ix.
A;Reference number: A34652; MUID:90147805; PMID:2302244
A;Accession: A34652
A;Status: preliminary
A;Molecule type: protein
A;Residues: 20-39 <YAZ>
C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protein
C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to c
C;Superfamily: calsequestrin
C;Keywords: calcium binding; cardiac muscle; glycoprotein; heart; skeletal muscle
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-406/Product: calsequestrin, cardiac and skeletal muscle #status experimental <NAT>
F;335/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 75.9%; Score 44; DB 1; Length 406;
Best Local Similarity 87.5%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDDPGF 8
Best Local Similarity 75.9%; Score 44; DB 2; Length 945;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 275 EDDDDPGF 282

RESULT 2

JN0901
endopeptidase Clp ATP-binding chain C - Arabidopsis thaliana
N;Alternate names: ATP-dependent Clp proteinase regulatory chain; ERD1 protein
N;Contains: adenosinetriphosphatase (EC 3.6.1.3)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: JN0901
R;Kiyosue, T.; Yamaguchi-Shinozaki, K.; Shinozaki, K.
Biochem. Biophys. Res. Commun. 196, 1214-1220, 1993
A;Title: Characterization of cDNA for a dehydration-inducible gene that encodes a Clp A.
A;Reference number: JN0901; MUID:94071876; PMID:7504470
A;Accession: JN0901
A;Molecule type: mRNA
A;Residues: 1-945 <XIY>
A;Cross-references: UNIPROT:P42762; GB:D17582; PIDN:EAA04506.1; PID:g497629
A;Note: this protein is homologous to the ATP-binding subunit of ATP-dependent Clp prote
C;Comment: This protein contains a putative chloroplast targeting signal at the amino-te
C;Comment: This protein interacts with a Clp-like protein and functions in the degradati
C;Genetics:
A;Gene: ERD1
C;Function: allows clpp to hydrolyze polypeptides and proteins, probably by a chaper
e activity; ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller
C;Superfamily: endopeptidase Clp ATP-binding chain
C;Keywords: ATP; GTP binding; hydrolase; molecular chaperone; nucleotide binding; P-loop
F;316-323/Region: nucleotide-binding motif A (P-loop)
F;383-388/Region: nucleotide-binding motif B
F;564-671/Region: nucleotide-binding motif A (P-loop)
F;732-737/Region: nucleotide-binding motif B
F;322/Binding site: ATP (Lys) #status predicted
F;670/Binding site: ATP (Lys) #status predicted

Query Match 75.9%; Score 44; DB 2; Length 945;
Best Local Similarity 80.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
Best Local Similarity 80.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 118 EDRDQPGFLG 127

RESULT 3

S22418
calsequestrin precursor, skeletal muscle - edible frog
N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-
C;Species: Rana esculenta (edible frog)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: S22418
R;Reves, S.; Vilsen, B.; Chiozzi, P.; Andersen, J.P.; Zorzato, F.
Biochem. J. 283, 767-772, 1992
A;Title: Molecular cloning, functional expression and tissue distribution of the cDNA en
A;Reference number: S22418; MUID:92272676; PMID:1375450
A;Accession: S22418
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-420 <TRE>
A;Cross-references: UNIPROT:P31231; EMBL:X64324; NID:g64276; PIDN:CAA45609.1; PID:g64277
C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protei
C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to c
C;Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in vi
C;Superfamily: calsequestrin
C;Keywords: calcium binding; glycoprotein; skeletal muscle
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-28/Domain: propeptide #status predicted <PRO>
F;23-420/Product: calsequestrin, fast skeletal muscle #status predicted <MAT>
F;338/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.7%; Score 41; DB 1; Length 420;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDDPGF 8
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 278 EDDDDPGF 285

RESULT 4

E84347
glycerol kinase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84347
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: E84347
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-510 <STO>
A;Cross-references: UNIPROT:Q9HNS5; GB:AE004437; NID:g10581402; PIDN:AAG20145.1; GSPDB:G1
C;Genetics:
A;Gene: GlpK
C;Superfamily: xylulokinase

Query Match 70.7%; Score 41; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DPDQFLG 10
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 238 DPDQFLG 244

RESULT 5

D71260
hypothetical protein TP0969 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: D71260
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 373-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: D71260
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-539 <COL>
A;Cross-references: UNIPROT:O83935; GB:AE001264; GB:AE000520; NID:g3323278; PIDN:AAC6592;
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0969

Query Match 67.2%; Score 39; DB 2; Length 539;
Best Local Similarity 77.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDDDDPGFL 9
Best Local Similarity 77.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 274 EDDDDPGFL 282

RESULT 6

G75608

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51476
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: 225394
A;Accession: T51476
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-526 <SAT>
A;Cross-references: UNIPROT:Q9LF52; EMBL:AL391150
A;Experimental source: cultivar Columbia; BAC clone K3M16
C;Genetics:
A;Map position: 5
A;Introns: 135/3; 218/2; 327/2; 393/1
A;Note: K3M16_70

Query Match 65.5%; Score 38; DB 2; Length 526;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDQDPDGFLG 10
:|||||
Db 163 NDDPDAPFG 171

RESULT 9
D84386
ATP-dependent RNA helicase homolog eIF-4A [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84386
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-784 <STO>
A;Cross-references: UNIPROT:Q9HWM5; GB:AE004437; NID:gi0581766; PIDN:AAG20456.1; GSPDB:GN
C;Genetics:
A;Gene: eif4a
C;Superfamily: ATP-dependent RNA helicase eIF-4A

Query Match 65.5%; Score 38; DB 2; Length 784;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDPDGFLG 10
:|||||
Db 274 DDDSDGQYG 283

RESULT 10
T49425
hypothetical protein B17C10.60 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49425
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: 225022
A;Accession: T49425
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1047 <SCH>
A;Cross-references: UNIPROT:Q9PE63; EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.60
A;Experimental source: BAC clone B17C10; strain OR74A
C;Genetics:

```

A:Gene: NCSP:B17C10.60
A:Map position: 6
A:Introns: 42/1; 55/1; 94/2; 126/3; 336/1; 489/2
C:Superfamily: Saccharomyces cerevisiae hypothetical protein YMR196w

Query Match      65.5%; Score 38; DB 2; Length 1047;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDPDGF 9
Db 183 EDDDPDELL 191

RESULT 11
AH3491
hypothetical cytosolic protein BMEI1918 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AH3491
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3491
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <KUR>
A:Cross-references: UNIPROT:Q8G3C5; GB:AE008917; PIDN:AAL53099.1; PID:g1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1918
A:Map position: I

Query Match      63.8%; Score 37; DB 2; Length 127;
Best Local Similarity 70.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDDDPDGF 10
Db 108 EDDVSGILG 117

RESULT 12
AD3306
periplasmic dipeptide transport protein precursor BMEI0434 [imported] - Brucella meliten
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AD3306
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3306
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <KUR>
A:Cross-references: UNIPROT:Q8YI10; GB:AE008917; PIDN:AAL51615.1; PID:g17982341; GSPDB:G
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0434
A:Map position: I

Query Match      63.8%; Score 37; DB 2; Length 186;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDPDGF 10
Db 87 DNGDPDNFLG 96

```

```

RESULT 13
AG3480
25K outer-membrane immunogenic protein precursor [imported] - Brucella melitensis (strain
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AG3480
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3480
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-228 <KUR>
A:Cross-references: UNIPROT:Q8YEP9; UNIPROT:Q8G333; GB:AE008917; PIDN:AAL53010.1; PID:g1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1829
A:Map position: I

Query Match      63.8%; Score 37; DB 2; Length 228;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDPDGFLG 10
Db 69 DIKPDGFLG 77

RESULT 14
H90015
hypothetical protein SA1999 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: H90015
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Uli, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H90015
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-243 <KUR>
A:Cross-references: UNIPROT:Q99S72; GB:BA000018; PID:g13701997; PIDN:BAB43289.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
C:Superfamily: conserved hypothetical protein b1120

Query Match      63.8%; Score 37; DB 2; Length 243;
Best Local Similarity 62.5%; Pred. No. 72;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DDDPDGFL 9
Db 62 EDDPEGFI 69

RESULT 15
S74913
sensory transduction system regulatory protein slr1400 - Synechocystis sp. (strain PCC 68
N:Alternate names: protein slr1400
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S74913
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996

```

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 s.
 A;Reference number: S74322; MUID:97061201; PMID:8905231
 A;Accession: S74913
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-380 <KAN>
 A;Cross-references: UNIPROT:P72936; EMBL:D90902; GS:AB001339; NID:g1652027; PIDN:BAA1695
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C;Superfamily: sensory transduction system regulatory protein; response regulator homolog
 C;Keywords: phosphoprotein
 F;6-117/Domain: response regulator homology <RRH>
 F;54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 63.8%; Score 37; DB 1; Length 380;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDDPDGF 8
 |||||
 Db 10 DDDPDNF 16

Search completed: September 20, 2005, 12:43:23
 Job time : 2.12607 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 5.07956 Seconds
(without alignments)
1008.117 Million cell updates/sec

Title: US-10-671-242-22
Perfect score: 58
Sequence: 1 EDDDPDGLG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	58	100.0	236	2	Q96IM4	Q96im4 homo sapien
2	58	100.0	285	2	Q6P0R3	Q6p0r3 homo sapien
3	58	100.0	295	2	Q6PIS7	Q6pis7 homo sapien
4	58	100.0	538	2	Q6SPF0	Q6spf0 homo sapien
5	54	93.1	550	2	Q6SPE9	Q6spe9 oryctolagus
6	46	79.3	872	2	Q6CYD6	Q6cyd6 kluyveromyc
7	45	77.6	509	2	Q75DH4	P19204 gallus gall
8	44	75.9	406	1	CAQ1 CHICK	Q75dh4 ashbya goss
9	44	75.9	640	2	Q94C10	P19204 gallus gall
10	44	75.9	945	1	ERD1 ARATH	P42762 arabidopsis
11	43	74.1	340	2	Q81XQ4	Q81xq4 homo sapien
12	43	74.1	346	2	Q8K2V8	Q8k2v8 mus musculu
13	43	74.1	351	2	Q9C0F9	Q9c0f9 homo sapien
14	43	74.1	364	2	Q69ZC8	Q69zc8 mus musculu
15	43	74.1	804	2	Q9VWD3	Q9vwd3 drosophila
16	41	70.7	273	2	Q6A9M3	Q6a9m3 propionibac
17	41	70.7	293	2	Q64RB7	Q64rb7 bacteroides
18	41	70.7	300	2	Q6PBF1	Q6pbf1 xenopus tro
19	41	70.7	354	2	Q96EM0	Q96em0 homo sapien
20	41	70.7	354	2	Q96LJ5	Q96lj5 homo sapien
21	41	70.7	396	2	Q6TU32	Q6tu32 aplysia cal
22	41	70.7	408	2	Q6AR41	Q6ar41 desulfotale
23	41	70.7	420	1	CAQ1_RANES	F31231 rana esculu
24	41	70.7	428	2	Q7ZWM0	Q7zwm0 xenopus lae
25	41	70.7	435	2	Q6ZTV1	Q6ztv1 homo sapien
26	41	70.7	450	2	Q7F9F0	Q7f9f0 oryza sativ
27	41	70.7	510	1	GLPK HALN1	Q9hns5 halobacteri
28	40	69.0	265	2	Q6WL19	Q6wl19 rhizobium s
29	40	69.0	316	2	Q82GF4	Q82gf4 streptomyce
30	40	69.0	324	2	Q9KXV9	Q9kxv9 streptomyce
31	40	69.0	409	2	Q6DI16	Q6di16 brachydanio

32 40 69.0 623 2 Q7PUC7
33 40 69.0 673 2 Q653C5
34 40 69.0 729 2 Q9GSZ5
35 40 69.0 859 2 Q6CL30
36 40 69.0 1167 2 Q8T2I8
37 40 69.0 1311 2 Q9VXJ5
38 39 67.2 77 2 Q8W1M3
39 39 67.2 84 2 Q8LP75
40 39 67.2 89 2 Q8LP76
41 39 67.2 90 2 Q8LP77
42 39 67.2 96 2 Q8LP79
43 39 67.2 99 2 Q8W1M4
44 39 67.2 108 2 Q83148
45 39 67.2 254 2 Q7NCD0

ALIGNMENTS

RESULT 1
Q96IM4 PRELIMINARY; PRT; 236 AA.
AC Q96IM4
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE LOC90378 protein (Hypothetical protein) (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007384; AAH07384.2; -;
DR EMBL; BC080588; AAH080588.1; -;
DR HSSP; P39769; 1KW4
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.

Q7puc7 anopheles g
Q653c5 oryza sativ
Q9gsz5 drosophila
Q6cl30 kluyveromyc
Q8t2i8 dictyosteli
Q9vxj5 drosophila
Q8w1m3 solanum cha
Q8lp75 lycopersico
Q8lp76 lycopersico
Q8lp77 lycopersico
Q8lp79 lycopersico
Q8w1m4 solanum cha
Q83148 murid herpe
Q7ncd0 gloeobacter

```

DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 58; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 227 EDDDPDGFGLG 236

RESULT 2
Q6P0R3 PRELIMINARY; PRT; 285 AA.
ID Q6P0R3
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 2
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 2
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 276 EDDDPDGFGLG 285

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
ID Q6PIS7
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
DE LOC90378 protein (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 2
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 286 EDDDPDGFGLG 295

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
ID Q6SPF0
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 58; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db |||||
529 EDDDPDGFLG 538

RESULT 5
Q6SPE9 PRELIMINARY; PRT; 550 AA.
AC Q6SPE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atherin.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453841; AAR24088.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM DOMAIN; 1.
SQ SEQUENCE 550 AA; 57019 MW; AF1CF2B780D879A7 CRC64;

Query Match 93.18; Score 54; DB 2; Length 550;
Best Local Similarity 90.0%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db |||||
541 EDDDPDGFLG 550

RESULT 6
Q6CYD6 PRELIMINARY; PRT; 872 AA.
AC Q6CYD6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P53076 Saccharomyces cerevisiae YGL227w VID30
DE singleton.
GN ORFNames=KLLA0A01221g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

```

```

RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolski M., Ozas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RN Nature 430:35-44(2004).
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382121; CAH02641.1; -.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR InterPro; IPR006595; CTLLH_C.
DR InterPro; IPR006594; LISH_C.
DR InterPro; IPR00342; RGS.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00622; SPRY; 1.
DR SMART; SM00449; SPRY; 1.
DR PROSITE; PS0897; CTLLH; 1.
DR PROSITE; PS0896; LISH; 1.
DR PROSITE; PS0896; LISH; 1.
SQ SEQUENCE 872 AA; 99640 MW; 8FACE92753B2DC8 CRC64;

Query Match 79.3%; Score 46; DB 2; Length 872;
Best Local Similarity 88.9%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDDDPDGFL 9
Db |||||
589 EDDDPDGFL 597

RESULT 7
Q75DH4 PRELIMINARY; PRT; 509 AA.
AC Q75DH4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ABR052Wp.
GN ORFNames=ABR052W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegeli S.E., Dietrich P.S., Brachat S., Lerch A., Gaffney T.,
RA Philippe P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016815; AAS50822.1; -.
DR AGD; ABR052W; -.
DR GO; GO:0016742; F:hydroxymethyl-, formyl- and related transfe. .; IEA.
DR GO; GO:0009058; P:ribosynthesis; IEA.
DR InterPro; IPR002376; Formyl transf. N.
SQ SEQUENCE 509 AA; 56603 MW; A66BAC28DD62FF88 CRC64;

Query Match 77.6%; Score 45; DB 2; Length 509;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFL 8
Db |||||

```

Db 302 EDDDPDGY 309

RESULT 8

CAQ1_CHICK STANDARD; PRT; 406 AA.

AC P19204;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DE 03-JUL-2004 (Rel. 44, Last annotation update)

DE Calsequestrin, skeletal muscle isoform precursor (Calsequestrin 1) (Aspartactin) (Laminin-binding protein).

GN Name=CAQ1;

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCB1_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=White leghorn;

RA Clegg D.O., Helder J.C., Hann B.C., Hall D.E., Reichardt L.F.;

RT "Amino acid sequence and distribution of mRNA encoding a major skeletal muscle laminin binding protein: an extracellular matrix-associated protein with an unusual COOH-terminal polyaspartate domain.";

RT J. Cell Biol. 107:699-705 (1988).

RN [2]

RP SEQUENCE OF 20-406 FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE=90358803; PubMed=2390076;

RA Yazaki P.J., Salvatori S., Dahms A.S.;

RT "Amino acid sequence of chicken calsequestrin deduced from cDNA: comparison of calsequestrin and aspartactin.";

RL Biochem. Biophys. Res. Commun. 170:1089-1095 (1990).

RN [3]

RP ERRATUM.

RX MEDLINE=91083667; PubMed=2260982;

RA Yazaki P.J., Salvatori S., Dahms A.S.;

RL Biochem. Biophys. Res. Commun. 173:763-763 (1990).

RN [4]

RP SEQUENCE OF 20-39.

RC STRAIN=White leghorn; TISSUE=Skeletal muscle;

RX MEDLINE=88331073; PubMed=3417768; DOI=10.1083/jcb.107.2.687;

RA Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;

RT "Isolation and characterization of a laminin-binding protein from rat and chick muscle.";

RL J. Cell Biol. 107:687-697 (1988).

RN [5]

RP SEQUENCE OF 20-39.

RC STRAIN=New Hampshire;

RX MEDLINE=90147805; PubMed=2302244;

RA Yazaki P.J., Salvatori S., Sabbadini R.A., Dahms A.S.;

RT "Calsequestrin, an intracellular calcium-binding protein of skeletal muscle sarcoplasmic reticulum, is homologous to aspartactin, a putative laminin-binding protein of the extracellular matrix.";

RL Biochem. Biophys. Res. Commun. 166:898-903 (1990).

CC -1- FUNCTION: Calsequestrin is a high-capacity, moderate affinity, calcium-binding protein and thus acts as an internal calcium store in muscle. The release of calcium bound to calsequestrin through a calcium release channel triggers muscle contraction. Binds 40 to 50 moles of calcium. Also binds laminin.

CC -1- SUBCELLULAR LOCATION: This isoform of calsequestrin occurs in the sarcoplasmic reticulum's terminal cisternae luminal spaces of fast skeletal muscle cells. Aspartactin is found in the basal lamina surrounding individual muscle fibers.

CC -1- TISSUE SPECIFICITY: Skeletal and heart muscle.

CC -1- SIMILARITY: Belongs to the calsequestrin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Y00789; CAA68743.1; -

DR EMBL; M58048; AAA48674.1; -

DR PIR; A31050; A31050.

DR HSP; P07221; LA8Y.

DR InterPro; IPR001393; Calsequestrin.

DR Pfam; PF01216; Calsequestrin; 1.

DR PRINTS; PR00312; CALSEQUESTRN.

DR PROSITE; PS00863; CALSEQUESTRN_1; 1.

DR PROSITE; PS00864; CALSEQUESTRN_2; 1.

KW Calcium-binding; Direct protein sequencing; Glycoprotein;

KW Muscle protein; Signal.

FT SIGNAL 1 19

FT CHAIN 20 406 Calsequestrin, skeletal muscle isoform.

FT DOMAIN 374 406 Poly-Asp.

FT CARBOHYD 335 335 N-linked (GlcNAc...) (Probable).

FT CONFLICT 148 148 V -> I (in Ref. 4).

SQ SEQUENCE 406 AA; 47151 MW; 928CF5A785A2DC7C CRC64;

Query Match 75.9%; Score 44; DB 1; Length 406;

Best Local Similarity 87.5%; Pred. No. 65;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGF 8

DB 275 EDDDPDGF 282

RESULT 9

Q94C10 PRELIMINARY; PRT; 640 AA.

ID Q94C10;

AC Q94C10;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE A1951070/K3K7_27.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P., Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY037264; AAK59865.1; -

DR HSP; Q9RA63; IQVR.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:000166; F:nucleotide binding; IEA.

DR GO; GO:0005515; F:protein binding; IEA.

DR GO; GO:0019538; P:protein metabolism; IEA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003959; AAA_ATPase_cent.

DR InterPro; IPR001270; Chaprinin_clp/B.

DR Pfam; PF00004; AAA; 1.

DR Pfam; PF02861; Clp_N; 2.

DR SMART; SM00382; AAA; 1.

DR PROSITE; PS00870; CLPAB_1; UNKNOWN_1.

KW ATP-binding.

SQ SEQUENCE 640 AA; 69610 MW; 4C477ECA608E84BB CRC64;

Query Match 75.9%; Score 44; DB 2; Length 640;

Best Local Similarity 80.0%; Pred. No. 1e+02; Mismatches 0; Indels 2; Gaps 0;
Matches 8; Conservative 0;

QY 1 EDDDDPGFLG 10
Db 118 EDRDPQFLG 127

RESULT 10

ERD1_ARATH STANDARD; PRT; 945 AA.

AC P42762;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE ERD1 protein, chloroplast precursor.
GN Name=ERD1; OrderedLocNames=At5g1070; ORFNames=K3K7.27;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94071876; PubMed=7504470;
RA Kiyose T., Yamaguchi-Shinozaki K., Shinozaki K.;
RT "Characterization of cDNA for a dehydration-inducible gene that
RT encodes a Clp A, B-like protein in Arabidopsis thaliana L.";
RL Biochem. Biophys. Res. Commun. 196:1214-1220(1993).
RN [2]

SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones";
RL DNA Res. 7:31-63(2000).
CC -I- FUNCTION: May interact with a clpp-like protease involved in
CC degradation of denatured proteins in the chloroplast.
CC -I- SUBCELLULAR LOCATION: Chloroplast (Potential).
CC -I- INDUCTION: By dehydration stress. Induced after one hour of
CC dehydration-stress and reaches maximal levels after 10 hours.
CC -I- SIMILARITY: Belongs to the clpA/clpB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; D17582; BAA04506.1; --
CC ENBL; AB023044; BAB10330.1; --
CC ENBL; AB017063; BAB10330.1; JOINED.
CC FR; JN0901; JN0901.
CC HSP; P03815; IJBK.
CC GeneFarm; 2731; 97.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC InterPro; IPR001270; Chaprinin_clpA/B.
CC InterPro; IPR004176; Clp_N.
CC Pfam; PF00004; AAA; 1.
CC PRINTS; PR00300; CLP_PROTASEA.
CC SMART; SM00382; AAA; 2.
CC PROSITE; PS00870; CLPAB_1; 1.
CC PROSITE; PS00871; CLPAB_2; 1.
KW ATP-binding; Chaperone; Chloroplast; Repeat; Transit peptide.
FT TRANSIT 1 ? Chloroplast (potential).

FT CHAIN ? 945 ERD1 protein.
FT DOMAIN 271 523 I.
FT DOMAIN 590 781 II.
FT NP_BIND 316 323 ATP (Potential).
FT NP_BIND 664 671 ATP (Potential).
SQ SEQUENCE 945 AA; 103234 MW; 81EF232C78F656B CRC64;

Query Match 75.9%; Score 44; DB 1; Length 945;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
Db 118 EDRDPQFLG 127

RESULT 11

Q81XQ4 PRELIMINARY; PRT; 340 AA.

AC Q81XQ4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KIAA1704.
GN Name=KIAA1704;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

SEQUENCE FROM N.A.

RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC039586; AAH39586.1; --
SQ SEQUENCE 340 AA; 38141 MW; F752A6A2917132E6 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 340;
Best Local Similarity 70.0%; Pred. No. 78;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
Db 83 DDDDDGDFG 92

RESULT 12

Q8K2V8 PRELIMINARY; PRT; 346 AA.

AC Q8K2V8;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DT 23-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE RIKEN cDNA 120001118 (Mus musculus 7 days embryo whole body cDNA,
 DE RIKEN full-length enriched library, clone:C430016019 product:similar
 DE to PROTEIN PHOSPHATASE 4 REGULATORY SUBUNIT 2).
 GN Name:120001118Rik;
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [7]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito D., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki R., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC029731; AAH29731.1; -;
 DR EMBL; AK082912; BAC38684.1; -;
 DR MGD; MGI:1914717; 120001118Rik.
 SQ SEQUENCE 346 AA; 38954 MW; F7EBC0CCD8267633 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 346;
 Best Local Similarity 70.0%; Pred. No. 80;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
 Db 92 DDDDDGFFG 91

RESULT 13

Q9COF9 PRELIMINARY; PRT; 351 AA.
 AC Q9COF9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE KIAA1704 protein (Fragment).
 GN Name:KIAA1704;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 DR EMBL; AB051491; BAB21795.1; -;
 FT NON TER 1
 SQ SEQUENCE 351 AA; 39406 MW; 6F5252CB5A6FE2F2 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 351;
 Best Local Similarity 70.0%; Pred. No. 81;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10
 Db 92 DDDDDGFFG 91

```

Db          94 DDDDDGFFG 103
RESULT 14
Q69ZC8      PRELIMINARY;          PRT;      364 AA.
AC Q69ZC8;
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE MKIAA1704 protein (Fragment).
GN Name=MKIAA1704;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AK173238; BAD32516.1; --
FT NON TER 1 1
SQ SEQUENCE 364 AA; 40946 MW; C9C3A38CBEB8D9A5 CRC64;

Query Match      74.1%; Score 43; DB 2; Length 364;
Best Local Similarity 70.0%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY          1 EDDDDGFFG 10
DB          100 DDDDDGFFG 109

RESULT 15
Q9VWD3      PRELIMINARY;          PRT;      804 AA.
AC Q9VWD3;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TEMBLrel. 26, Last annotation update)
DE CG12531-PA.
GN ORFNames=CG12531;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Sutton G.G., Wortman J.R., Venter L., White O., Adams M.D.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abrial J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Caciulea E., Center A., Chandra I.,
RA Chert J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupry J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003513; AAF49012.1; --
DR FlyBase; FBN0031064; CG12531.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005279; P:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA_rel_permease.
DR InterPro; IPR004841; Permease_region.
DR Pfam; PF00324; AA_permease; 2.
DR Transmembrane; Transport.
KW

```

SQ SEQUENCE 804 AA; 86555 MW; 6D13A1DB1D2EC1AB CRC64;
Query Match 74.1%; Score 43; DB 2; Length 804;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 DDDPDGFLG 10
| | | | |
| | | | |
Db 483 DDSFEGFLG 491

Search completed: September 20, 2005, 12:41:47
Job time : 6.07956 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 5.5202 Seconds
(without alignments)
733.538 Million cell updates/sec

Title: US-10-671-242-22
Perfect score: 58
Sequence: 1 EDDPDGFLG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	10	9	US-09-962-055-22
2	58	100.0	10	9	US-09-976-740-22
3	58	100.0	10	13	US-10-023-529-22
4	58	100.0	10	13	US-10-023-523-22
5	58	100.0	10	15	US-10-616-187-22
6	58	100.0	10	15	US-10-671-242-22
7	58	100.0	217	9	US-09-962-055-7
8	58	100.0	217	9	US-09-976-740-7
9	58	100.0	217	13	US-10-023-529-7
10	58	100.0	217	13	US-10-023-523-7
11	58	100.0	217	15	US-10-616-187-7

58	100.0	217	15	US-10-671-242-7	Sequence 7, Appli
58	100.0	241	9	US-09-925-298-665	Sequence 665, App
58	100.0	241	14	US-10-102-806-665	Sequence 665, App
58	100.0	538	9	US-09-976-740-43	Sequence 43, Appl
58	100.0	538	13	US-10-023-529-43	Sequence 43, Appl
58	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
58	100.0	538	15	US-10-616-187-43	Sequence 43, Appl
58	100.0	538	15	US-10-671-242-43	Sequence 43, Appl
54	93.1	232	9	US-09-962-055-3	Sequence 3, Appli
54	93.1	232	9	US-09-976-740-3	Sequence 3, Appli
54	93.1	232	13	US-10-023-529-3	Sequence 3, Appli
54	93.1	232	13	US-10-023-523-3	Sequence 3, Appli
54	93.1	232	15	US-10-616-187-3	Sequence 3, Appli
54	93.1	232	15	US-10-671-242-3	Sequence 3, Appli
54	93.1	252	9	US-09-962-055-4	Sequence 4, Appli
54	93.1	252	9	US-09-976-740-4	Sequence 4, Appli
54	93.1	252	13	US-10-023-529-4	Sequence 4, Appli
54	93.1	252	13	US-10-023-523-4	Sequence 4, Appli
54	93.1	252	15	US-10-616-187-4	Sequence 4, Appli
54	93.1	252	15	US-10-671-242-4	Sequence 4, Appli
54	93.1	317	9	US-09-962-055-2	Sequence 2, Appli
54	93.1	317	9	US-09-976-740-2	Sequence 2, Appli
54	93.1	317	13	US-10-023-529-2	Sequence 2, Appli
54	93.1	317	13	US-10-023-523-2	Sequence 2, Appli
54	93.1	317	15	US-10-616-187-2	Sequence 2, Appli
54	93.1	317	15	US-10-671-242-2	Sequence 2, Appli
54	93.1	550	9	US-09-976-740-47	Sequence 47, Appl
54	93.1	550	13	US-10-023-529-47	Sequence 47, Appl
54	93.1	550	13	US-10-023-523-47	Sequence 47, Appl
54	93.1	550	15	US-10-616-187-47	Sequence 47, Appl
54	93.1	550	15	US-10-671-242-47	Sequence 47, Appl
44	75.9	640	17	US-10-732-923-7094	Sequence 7094, Ap
44	75.9	945	9	US-09-812-350-3	Sequence 3, Appli
44	75.9	945	17	US-10-732-923-7090	Sequence 7090, Ap

ALIGNMENTS

RESULT 1
US-09-962-055-22
; Sequence 22, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-962-055-22

Query Match      100.0%; Score 58; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 2
US-09-976-740-22
; Sequence 22, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-22

Query Match      100.0%; Score 58; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 3
US-10-023-529-22
; Sequence 22, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-22

Query Match      100.0%; Score 58; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 4
US-10-023-523-22
; Sequence 22, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-22

Query Match      100.0%; Score 58; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-22

Query Match      100.0%; Score 58; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10
```

```
RESULT 4
US-10-023-523-22
; Sequence 22, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-22

Query Match      100.0%; Score 58; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10
```


RESULT 5
US-10-616-187-22
; Sequence 22, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-22

Query Match 100.0%; Score 58; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 1 EDDDPDGFGLG 10

RESULT 6
US-10-671-242-22
; Sequence 22, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-22

Query Match 100.0%; Score 58; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 1 EDDDPDGFGLG 10

RESULT 7
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match 100.0%; Score 58; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 208 EDDDPDGFGLG 217

RESULT 8
US-09-976-740-7


```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-7

Query Match      100.0%; Score 58; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 208 EDDDPDGFLG 217

RESULT 12
US-10-671-242-7
; Sequence 7, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 58; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 208 EDDDPDGFLG 217

RESULT 13
US-09-925-298-665
; Sequence 665, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 58; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 208 EDDDPDGFLG 217

RESULT 14
US-10-102-806-665
; Sequence 665, Application US/10102806
; Publication No. US2003005421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 665
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-665

Query Match      100.0%; Score 58; DB 9; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 232 EDDDPDGFLG 241

RESULT 15
US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
```

; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-43

Query Match 100.0%; Score 58; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
|||||
Db 529 EDDDPDGFGLG 538

Search completed: September 20, 2005, 12:53:07
Job time : 6.5202 secs